



Munich Personal RePEc Archive

## **Reddit's self-organised bull runs: Social contagion and asset prices**

Semenova, Valentina and Winkler, Julian

Institute of New Economic Thinking at the Oxford Martin School,  
Department of Mathematics, University of Oxford, Department of  
Economics, University of Oxford

5 May 2021

Online at <https://mpra.ub.uni-muenchen.de/107575/>  
MPRA Paper No. 107575, posted 08 May 2021 07:29 UTC

# Reddit's self-organised bull runs: Social contagion and asset prices

Valentina Semenova<sup>\*1,2</sup> and Julian Winkler<sup>+1,3</sup>

<sup>1</sup>*Institute of New Economic Thinking at the Oxford Martin School*

<sup>2</sup>*Department of Mathematics, University of Oxford*

<sup>3</sup>*Department of Economics, University of Oxford*

May 5, 2021<sup>‡</sup>

## Abstract

This paper develops an empirical and theoretical case for how ‘hype’ among retail investors can drive large asset fluctuations. We use the dataset of discussions on WallStreetBets (WSB), an online investor forum with over nine million followers as of April 2021, to show how excitement about trading opportunities can ripple through an investor community with large market impacts. This paper finds empirical evidence of psychological contagion among retail investors by exploiting differences in stock price fluctuations and discussion intensity. We show that asset discussions on WSB are self-perpetuating: an initial set of investors attracts a larger and larger group of excited followers. Sentiments about future stock performance also spread from one individual to the next, net of any fundamental price movements. Leveraging these findings, we develop a model for how social contagion impacts prices. The proposed model and simulations show that social contagion has a destabilizing effect on markets. Finally, we establish a causal relationship between WSB activity and financial markets using an instrumental variable approach.

JEL codes: D91, G14, G41.

## Acknowledgements

We thank Rick Van der Ploeg, J. Doyné Farmer, Xiaowen Dong and Steve Bond for supervising this research. We are grateful to Francis DiTraglia, Kevin Sheppard, Marteen Scholl, John Pougué-Biyong, Ilan Strauss, Jangho Yang, Torsten Heinrich, François Lafond, Matthias Winkler, José Moran, Farshad Ravasan, Mirta Galesic, Renaud Lambiotte and participants of the Royal Economic Society annual conference for their helpful comments and insightful questions. We thank Baillie Gifford and the Institute for New Economic Thinking at the Oxford Martin School for funding our work at the University of Oxford.

---

<sup>\*</sup>valentina.semenova@maths.ox.ac.uk

<sup>†</sup>julian.winkler@economics.ox.ac.uk

<sup>‡</sup>First version uploaded online January 20<sup>th</sup>, 2021

# 1 Introduction

A trite yet fundamental question in economics is: What causes large asset price fluctuations? A ten-fold rise in the price of GameStop equity, between the 22<sup>nd</sup> and 28<sup>th</sup> of January 2021, demonstrates that herding behaviour among retail investors is an important contributing factor. On this occasion, thousands of retail investors launched a speculative attack on the short positions held in GameStop by hedge funds, using social media as their coordination platform.

As academics and regulators alike grapple with the implications, many wonder whether large-scale coordination among retail investors is the new ‘modus operandi’, or a one-off fluke. We argue that this is a new manifestation of a well-established global phenomenon. Social media has changed the fabric of society. Polarization, the spread of fake news and other societal challenges are some of its documented consequences (Tucker et al. 2018). 4.2 billion people, or 53.6% of the world population, are active social media users, each just a few clicks away from the next popular phenomenon.<sup>1</sup> Now, a growing audience turns to social media for promising stock market gambles.

To the extent that humans are a social species, personal interactions are likely to play a role in financial decision-making. Consider, for example, that many bank runs (such as the Japanese financial crisis in 1927 or the Swedish bank run of 2011) have been triggered by a simple rumour. Studies of social dynamics in stock market activity date back to Shiller (1984), a seminal paper reminding us that ‘investing in speculative assets is a social activity.’ Internet communities are simply conduits, enabling retail investors to come together at an unprecedented scale.

Nevertheless, we lack a holistic theory, backed by empirical evidence, for how social interactions drive investment decisions. Practical difficulties and lack of data have restricted many researchers to controlled laboratory experiments, whose external validity remains unchecked. The new investing climate, enabled by communications technology, offers both an extraordinary opportunity and urgent need to understand the role that social dynamics play.

This paper sets out to reconcile the extreme nature of the behaviours on social media with economic theory. We use text data from the ‘WallStreetBets’ (WSB) Reddit forum to gain insight into investor social dynamics. Reddit is a social news aggregation site, ranked 19<sup>th</sup> most visited website in the world, as of April 2021.<sup>2</sup> Reddit, and WSB more specifically, have several desirable features for this line of research. The WSB forum is completely anonymous, with the quality of content alone driving further discussion. Users try to garner a following by galvanising their peers with their ballooning balances and precipitous losses, which they avidly post online. Finally, WSB has gained a colossal following: the forum currently has over nine million self-described ‘degenerates’.<sup>3</sup> Putting these numbers into perspective, The Times newspaper recently boasted 7.5 million subscriptions.<sup>4</sup>

We first document evidence of social *contagion* on the forum. The working hypothesis is that users’ expressed interest in assets depends on the discussions they engage with. Demand, much like an epidemic, ‘dies out’ if the asset is not discussed. Accurately evaluating contagion from one interested user to another is not straightforward due to endogeneity (Aral et al. 2009). This issue is addressed with a two-fold approach: i) a panel regression, controlling for outside factors, and ii) an opinion dynamics model that matches users on observable characteristics. The results demonstrate significant contagion in asset demand after accounting for endogeneity. Users who comment on discussions about an asset are between four and nine times more likely to subsequently start a new conversation about the asset themselves, compared to their matched counterparts in the control group.

The next question is whether investment strategies also transmit between users: in other words,

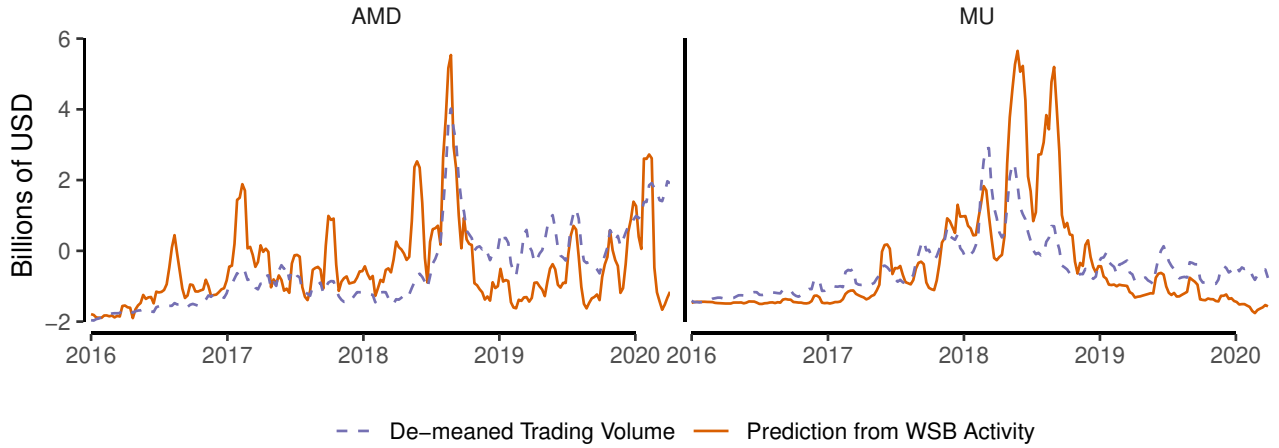


Figure 1: **Hype-Driven WSB Activity and Trading Volumes**; the predicted number of new users interested in a ticker, for a given week, correlates strongly with observed weekly trading volumes for AMD (*Advanced Micro Devices*) and MU (*Micron Technology*). Both series are plotted as four week rolling moving averages. See Section 4.4 for details.

whether we observe *consensus*. We measure ‘buy’ versus ‘sell’ strategies via the expressed sentiment about future asset performance in WSB submissions. The spread of sentiments about an asset can be measured by the associated spillover from the network of discussions. This network links submissions by their authors’ comments to older submissions mentioning the same asset. The results indicate that sentiments in submissions correlate strongly and significantly with the sentiments of neighbours, especially for bearish posts; the submission from a user who commented on a bearish post in the past is almost 50% more likely to be bearish than indifferent, net of market factors. Social interactions thus play an important role in determining investor sentiment. The asymmetry of the transmission mechanism of bullish versus bearish sentiment is of particular interest. Despite users encouraging risk-taking, there is strong evidence for risk-aversion and panic selling.

After characterizing the nature of social dynamics, a crucial question remains: what are the implications for financial market stability? Some may argue that the key dynamics, *contagion* and *consensus*, could lead to the fast spread of news and quick price convergence to an asset’s ‘true’ market value. Others may consider the potential for unchecked rumours to permeate the investor community, driving the market to diverge from fundamentals and to exhibit greater volatility.

In order to tackle this question, we propose a simple model with two mechanisms through which social dynamics impact asset prices: i) ‘contagion’, which generates the hump-shaped pattern in demand for an asset over time, as new investors initially buy into the hype but eventually become uninterested in an asset, and ii) ‘consensus’, which captures the extent of coordination in buying/selling amongst these investors. We borrow from Shiller (2017) to describe the rise and subsequent fall in investor interest. A discrete time simulation of our model, based on parameters initiated from WSB data, demonstrates that assets experience an initial, slow rise in value followed by a crash and period of high volatility.

We validate the existence of a causal link between WSB activity and the stock market through a Two-Stage-Least-Squares approach. We predict new users discussing an asset on WSB using historical WSB and market data: these new users are ‘hype’ investors, driven to join the trading activity by the excitement of their peers rather than by news or fundamentals. The predicted number of new

authors, plus the lag in average sentiment, are fitted to data on price returns, volatility, and trading volumes, after removing temporal and cross-sectional means. Our key finding is that our instrumental variable performs particularly well in explaining asset trading volumes. Figure 1 displays this result for two heavily discussed stocks, namely Advanced Micro Devices (AMD) and Micron Technology (MU). Hype investors appear cyclical: interest, built up over several weeks, eventually fades out. Their activity is correlated to negative asset returns and higher volatility. These findings offer empirical evidence of the fact that the bull runs generated by WSB activity are destabilising.

The next section comprehensively describes the data source and relevant variables. Section 3 brings to light empirical evidence of investor social dynamics. Section 4 presents a behavioural model for how asset interest and sentiments spread among investors, with implications for asset prices. It presents a causal relationship between WSB activity and market variables. Section 5 concludes.

## 2 What is *WallStreetBets*?

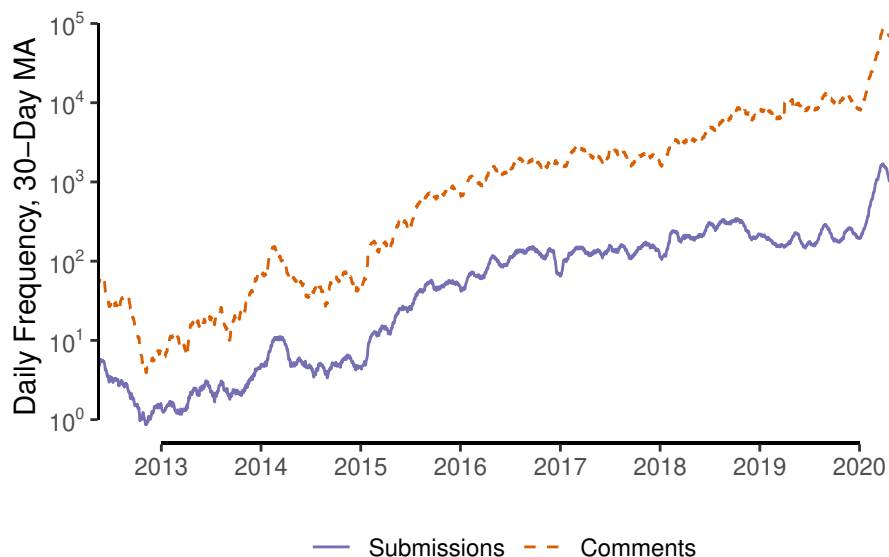
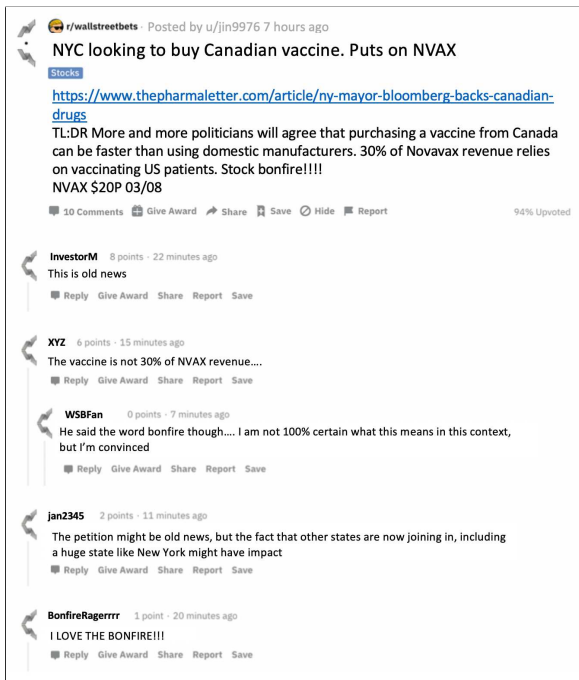


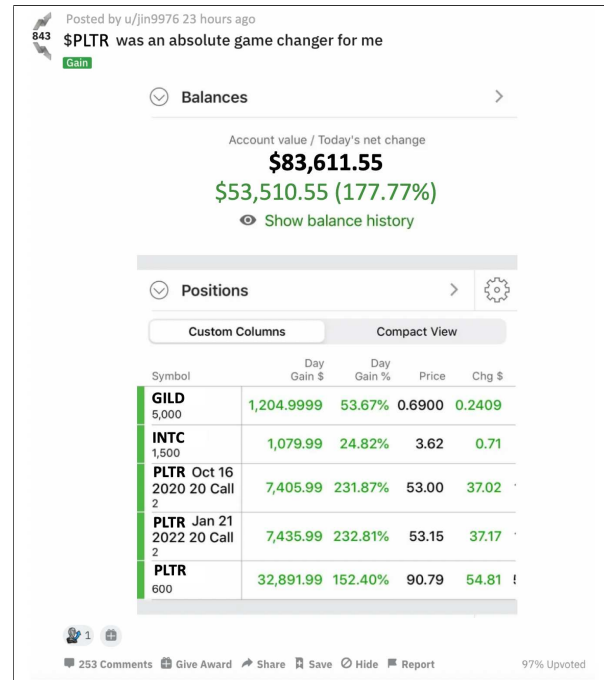
Figure 2: **Daily Activity on WSB Plotted on a Logarithmic Scale**; the daily submission and comment counts, averaged over 30 days, demonstrate a persistent exponential increase from 2015 to 2020, with a substantial jump in early 2020.

Reddit, launched in 2005, is a social news aggregation, web content rating, and discussion website. It was ranked as the 19<sup>th</sup> most visited site globally in April 2021<sup>5</sup>, with over 430 million anonymous users by the end of 2019<sup>6</sup>. The website’s contents are self-organized by subject into smaller sub-forums, ‘subreddits,’ to discuss a unique, central topic.

Within subreddits, users make titled posts (also called submissions), typically accompanied with a body of text or a link to an external website. These submissions can be commented and *upvoted* or *downvoted* by other users. A ranking algorithm raises the visibility of a submission with the amount of upvotes it receives, but lowers it with age. Therefore, the first posts that visitors see are i) highly upvoted, and ii) recent. Comments on a post are also visible, are subject to a similar scoring system, and can, themselves, be commented on.



(a) A Typical Discussion on WSB



(b) A Sample Screenshot of User Profits

**Figure 3: What Does WSB Look Like?** These snapshots exemplify typical discussions on WSB. The exact text, usernames, and conversation details have been modified to protect user identities.

The WSB subreddit was created on January 31, 2012, and reached one million followers in March 2020<sup>7</sup>. As per a Google survey from 2016, the majority of WSB users are ‘young, male, students that are inexperienced investors utilizing real money (not paper trading); most users have four figures in their trading account’<sup>8</sup>. The conversation guidelines outlined by the moderators of WSB handily demonstrate the financial focus and whimsical tone of discussions:<sup>9</sup>

- Discussion about day trading, stocks, options, futures, and anything market related,
- Charts and Technical Analysis,
- Shower before posting,
- Some irresponsible risk taking,
- People sharing trades, ideas, observations.

We gain insight into the content of WSB discussions through a biterm topic model (Yan et al. 2013). The model detects a mix of conversations, from those that follow specific real world trends, such as the China trade war or the COVID-19 pandemic, to those that hype positions on a handful of stocks and indices, most famously the GameStop (GME) short squeeze. Some topics persist across time: people consistently ask for advice about trading accounts and anonymously share details of how their trading is affecting their personal lives. On the other hand, topics concerned with specific economic interests wax and wane. Two examples of this are the uptick in submissions discussing GME and Robinhood account trading limits, coinciding with the GME short squeeze, and the COVID-19 topic, which is negligible until January 2020, but gains prominence in the subsequent months. A full description of the topic model and the temporal trends of topics are presented in Appendix A.8.

The subreddit’s size grew exponentially since 2015, plotted in Figure 2. Two jumps are notable: a smaller, seemingly idiosyncratic rise in early 2018, and a sharp spike during the COVID-19 pandemic. Figure 3a displays a typical exchange on the WSB forum: individuals discuss stock-related

news and their sentiments on whether this will affect stock prices in the future. In addition to market discussions, there is ample evidence of users pursuing the investment strategies encouraged in WSB conversations. Users post screenshots of their investment gains and losses, which moderators are encouraged to verify, as illustrated in Figure 3b. These observations are reminiscent of Shiller (2005) in his definition of an asset bubble as:

A situation in which news of price increases spurs investor enthusiasm which spreads by psychological contagion from person to person, in the process amplifying stories that might justify the price increases and bringing in a larger and larger class of investors, who, despite doubts about the real value of an investment, are drawn to it partly through envy of others' successes and partly through a **gambler's excitement**.

All posts made on Reddit, plus their metadata, can be queried via Reddit's API, as well as other sources. In what follows, we downloaded data on WSB using the PushShift API<sup>10</sup>. The only caveat of PushShift is that all data are recorded at the time of posting.

The full dataset consists of two parts. The first is a total of 452,720 submissions, with their authors, titles, text and timestamps. The second is comprised of 15.4 million comments, with their authors, text, timestamp, and their linked comment or post. The following sections will predominantly rely on submissions for text data, since they are substantially richer. Comments are largely used to trace user activity and, subsequently, the interaction between discussants. User submissions data, external to WSB, is also downloaded from PushShift, but is only available through April, 2020. The availability of external user data forces us to rely on data until April 30<sup>th</sup>, 2020 for our analysis, unless otherwise specified.

## 2.1 Identifying Asset Discussions

In order to understand how users discuss specific assets, we extract mentions of 'tickers' from the WSB submissions text data. A ticker is a short combination of capital letters, used to identify an asset in the financial markets. For example, 'AAPL' refers to shares in Apple, Inc. Appendix A.1 documents how tickers are extracted from submissions. Table 7 in Appendix A.1 displays the twenty tickers that feature most prominently in WSB conversations. These are typically shares in technology firms, such as AMD or FB. A handful of indices are also present, notably the S&P 500 (SPY) and a gold ETF (JNUG).

A small fraction of the 4,650 tickers we extract dominate the discourse on WSB. 90% of tickers are mentioned fewer than 31 times, and more than 60% are mentioned fewer than five times. The frequency distribution of tail of ticker mentions demonstrates this point, for which Figure 4 displays a QQ-plot. We arbitrarily selected tickers with the number of mentions in the top 10<sup>th</sup> percentile. Even though threshold of mentions for this top decile is 30 submissions, the most popular, SPY, features in almost 8,000 submissions. The orange crosses in Figure 4 locate the empirical densities, on a log scale, which are plotted against the theoretical quantiles of an exponential distribution on the x-axis. Under the assumption that ticker mentions are heavy-tailed (similarly to vocabulary distributions), the logarithm of the mentions follows an exponential distribution, with the intercept at the threshold, and the slope equal to the inverse of the tail index. Indeed, the linear fit is close to perfect, supporting the assumption that the popularity of assets in WSB is heavy-tailed, with an estimated tail exponent of approximately 1.03. In what follows, we used submissions for which we identified a single ticker, unless otherwise specified, forming a dataset of 103,205 submissions with unique ticker mentions by our cutoff date.

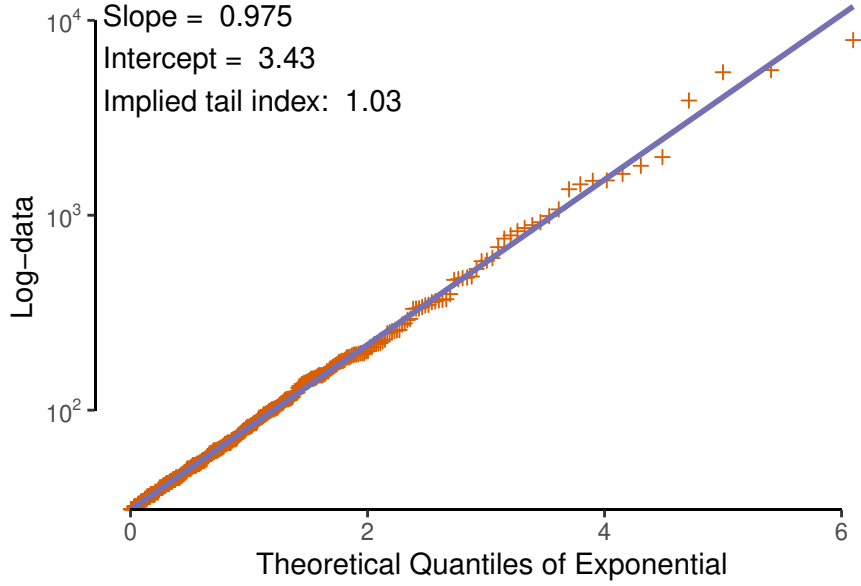


Figure 4: **QQ Plot of the Tail in Ticker Mentions on WSB**; the number of submissions for each ticker (on a log-scale) is plotted against the theoretical quantiles of an exponential distribution. Quantiles are calculated as  $q(i) = -\log(1 - i/(N + 1))$ , where  $N$  is the number of observations, and  $i$  the order of the statistic, from 1 to  $N$ . The linear fit suggests that the data follows a Pareto distribution, with the tail index equal to the inverse of the slope. The threshold for a ticker to be part of the ‘tail’ is 31 mentions; note the intercept, at  $\exp(3.43) \approx 31$ .

## 2.2 Sentiments about Assets

In order to thoroughly understand the social dynamics of asset discussions, it is not sufficient to simply identify what assets are being discussed, as in Section 2.1; it is important to understand what is being said about them. Our goal, with regards to the text data in WSB, is to gauge whether discussions on certain assets express an expectation for their future price to rise, the ‘bullish’ case, to fall, the ‘bearish’ case, or to remain unpredictable, the ‘neutral’ case.

A series of studies link sentiment, measured through diverse approaches, to stock market performance (Garcia 2013, Tetlock 2007, Bollen et al. 2011). Gentzkow et al. (2019) offer a thorough review. Many of these works use lexicon approaches, whereby documents are scored based on the prevalence of words associated with a certain sentiment. Recently, machine learning has offered alternative, powerful tools, such as Google’s Bidirectional Encoder Representations from Transformers (BERT) algorithm (Devlin et al. 2018). The BERT algorithm trains a final layer of nodes in a neural network from a pre-trained classifier on labelled data. The classifier itself is a neural net, pre-trained by Google on a corpus of Wikipedia entries to i) predict the probability distribution of words appearing in a given sentence (Masked Language Modeling), and ii) predict the relationship between sentences (Next Sentence Prediction).

Among other alternatives, we pursued a supervised-learning approach to identify the sentiment expressed about an asset within a WSB submission. This required a training dataset, for which we manually labelled 2,581 random submissions with unique ticker mentions as either ‘bullish,’ ‘bearish’ or ‘neutral,’ with respect to the authors’ expressed expectations for the future price. We used the BERT algorithm for labeling. Work not shown here implements an alternative regression-based approach as a robustness check, but BERT performs better out-of-sample. We discuss BERT’s



results and accuracy in Appendix A.2.

### 3 Social Dynamics among Retail Investors: Evidence of *Contagion* and *Consensus*

Do investors express demand for an asset because of an independently perceived payoff, or because of another investor’s stated interest? The endogenous dynamics in the WSB community can be estimated as a spillover rippling through a network of people, linked by their interactions on WSB.

Section 3.1 quantifies the spread in asset interest through the WSB network. It pursues a two-fold approach: i) a regression estimating the probability of current discussants of an asset to attract a greater following, and ii) matching on observables, outlined in Leng et al. (2018), Lehmann & Ahn (2018), in order to estimate a causal link between a user’s engagement with a stock and their future posting activity on that ticker. Both approaches attempt to mitigate the problems arising from homophily in endogenous link formation in networks, discussed in Aral et al. (2009).

Section 3.2 studies consensus among investors: if a user engages in a discussion about an asset with an expressed sentiment, what is the probability they adopt the same sentiment? The evolution of collective dynamics on social networks are studied in many contexts, including healthcare outcomes and product adoption, among others (Christakis & Fowler 2008, Lehmann & Ahn 2018). Their methodology builds a logistic regression model for the dependent variable at time  $t + 1$  as a function of demographic attributes and the status of the dependent variable among contacts at time  $t$ . This paper pursues a similar approach as it is a fruitful, simple way to gain an estimate of peer influence. The results demonstrate that the WSB network exhibits significant sentiment contagion, as older submissions influence future sentiments, to the degree that users interact.

#### 3.1 Contagion in Asset Demand

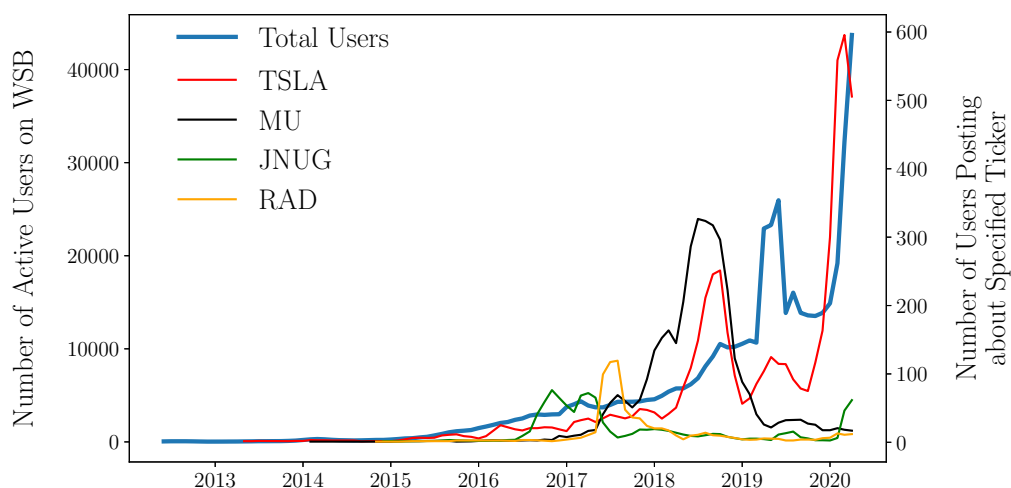


Figure 5: **Ticker Activity versus Overall WSB Activity**; the number of unique monthly active users (users who comment or post) on the left y-axis plotted against the unique number of posters about a specific ticker (on the right y-axis). Both are plotted as 3-month rolling means.

This section sets out to demonstrate empirically that asset interest, among investors on the WSB forum, is subject to contagion. Figure 5 shows how stock discussions permeate the forum for a

select subset of popular stocks: we plot the three-month rolling average of the number of unique authors who post about a specific ticker. Visually, an initial group of users appears to attract a larger following. Eventually, interest in a specific asset dissipates, as users turn to new opportunities. Such behaviour is reminiscent of a Kermack-McKendrick Susceptible-Infected-Recovered (SIR) epidemic model, proposed for the economic context in Shiller (2017).

First, we perform a least-squares regression analysis to isolate the impact that discussions on the forum have on begetting further discussions. We find that conversations on WSB indeed have an endogenous component, which lends credibility to the fact that social contagion plays a role in asset interest. Second, we perform a matching on observables exercise both to validate our initial results, and isolate a mechanism through which contagion occurs.

Table 1: Time Dynamics of Assets Discussed on WSB

|                          | <i>Dependent variable: <math>\Delta A_{i,t}</math></i> |                  |                  |
|--------------------------|--|------------------|------------------|
|                          | (1)  | (2)              | (3)              |
| $A_{i,t-1}$              | 0.40*** (0.002)  | 0.40*** (0.002)  | 0.38*** (0.002)  |
| $\sum_{j=1}^3 A_{i,t-j}$ | -0.06*** (0.001)                                       | -0.06*** (0.001) | -0.06*** (0.001) |
| $R_{i,t}^e$              |  |                  | -1.07** (0.47)   |
| $\sigma_{i,t}$           |  |                  | 11.64*** (0.34)  |
| $Q_{i,t}$                |  |                  | 0.48*** (0.01)   |
| $\alpha$                 | 0.15*** (0.01)   | -0.002 (0.01)    | -0.01 (0.01)     |
| Model                    | Pooled   | Within-week      | Within-week      |
| N                        | 118,024  | 118,024          | 116,326          |
| Adjusted R <sup>2</sup>  | 0.49   | 0.48             | 0.51             |

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

*Notes:* This table presents coefficients for an OLS model estimating the number of new users,  $\Delta A_{i,t}$ , discussing a ticker on WSB from the set of independent variables in Eq. 1:  $A_{i,t-1}$  is the number of unique authors who posted a submission on ticker  $i$  in weeks  $t-1$  to  $t-3$ , and  $\sum_{j=1}^3 A_{i,t-j}$  is the cumulative count of unique authors. Column (1) estimates the OLS coefficients by pooling all the available data, and column (2) subtracts the cross-sectional mean in a given week from each variable, to control for unobserved time heterogeneity. In particular, we may be concerned with WSB boasting a larger following in recent years. In column (3), we include controls for the average price return,  $R_{i,t}^e$ , price return standard deviation,  $\sigma_{i,t}$ , and average dollar trading volumes,  $Q_{i,t}$ , for asset  $i$  in week  $t$ . Details on the construction of financial variables are provided in Appendix A.3.

**A Simple Estimate for Contagion** Perhaps the simplest framework to test whether discussions are self-perpetuating is to look at whether a current user discussing an asset attracts new users to the conversation. The hypothesis is that the number of new users posting about a ticker at time  $t$  increases with the number of users discussing the ticker at the previous time step,  $t-1$ . The number of new users who join the discussion is expected to decline with the *cumulative* number of users who discussed the asset in the past since, at a certain point, users may decide that the investment opportunity expired, and seek fresh ideas.

The relationship between the current set of discussants of an asset and the entry of new users into the conversation may be driven by unobserved confounding factors. To assess this possibility,

we evaluate the effect using a panel framework, controlling for exogenous variables and temporal shifts. Our model of choice is an Ordinary Least Squares (OLS) regression:

$$\Delta A_{i,t} = \alpha + cA_{i,t-1} - r \sum_{j=1}^3 A_{i,t-j} + \beta' X_{i,t} + \varepsilon_{i,t}, \quad (1)$$

where  $\Delta A_{i,t}$  is a measure for the number of new users who post a submission on asset  $i$  within calendar week  $t$ ;  $A_{i,t-1}$  is the stock of existing users;  $\sum_{j=1}^T A_{i,t-j}$  is the cumulative stock of users;  $X_{i,t}$  is a vector of controls;  $\varepsilon_{i,t}$  is an error term. We find the number of existing users,  $A_{i,t-1}$ , by counting the number of unique authors who posted a submission on asset  $i$  between weeks  $t-3$  and  $t-1$ . We calculate the cumulative stock of users,  $\sum_{j=1}^T A_{i,t-j}$ , by summing  $A_{i,t-j}$ s, from an assumed starting point  $j=3$  to  $j=1$ . Our cumulative stock of users proxies the total amount of coverage the ticker has received in the last three weeks; we expect the number of new users to decline as this cumulative stock increases. We experiment with different time horizons and observe that our results remain largely unchanged; we choose to use three weeks of data to predict activity in the following week based on the observation that ticker discussions evolve over several months, as shown in Figure 5. In order to get reasonable coverage, we restrict ourselves to tickers mentioned over 30 times, coinciding with the tail tickers in Figure 4.

The key results are displayed in Table 1, estimating the coefficients from Eq. 1. We observe that enthusiasm about a ticker spreads from one user to the next. This effect is reflected by the estimate for coefficient  $c$ . The estimated coefficient  $r$ , on the other hand, proxies for the fact that, eventually, the set of new users who could become interested in a given asset is depleted, while current discussants move onto the next exciting stock.  $c$  and  $r$  can be likened to the ‘contagion rate’ and ‘recovery rate’ within the epidemiological SIR model: as more people get infected, they will infect others (captured by positive  $c$ ), however, as individuals recover, the population becomes immune (captured by a negative  $r$ ). The coefficients are statistically significant across all variants of our model. We also find an adjusted  $R^2$  value of approximately 0.5 for all constructions of our model. Almost half of the variance in the number of new authors is explained by our two contagion variables.

Let us consider the implications of the most basic construction of the model, presented in column (1). If ten authors posted a submission on asset  $i$  in the preceding week ( $t-1$ ), the model estimates that four new authors post their own submission on the same asset in the subsequent week ( $t$ ). Now consider a similar scenario, differing only in the fact that 50 different authors post about the asset two weeks before (at  $t-2$ ). Our model would now estimate that 18 new authors emerge: calculated as  $0.40 \times 60 - 0.06 \times 110 + 0.15$ . In the latter case, the model captures that, despite the high number of authors currently involved in the ticker discussion, the conversation is potentially on a decline.

Column (2) addresses an important problem concerning the assumption of a fixed population of users. Clearly, the number of users, seen in Figure 2, is anything but fixed. To account for this time pattern, we calculate the within-week deviations of all variables, by subtracting their cross-sectional averages in each week  $t$ . This procedure is identical to introducing week dummies, or time fixed effects. Column (2) demonstrates that the coefficients are entirely robust, with the exception of the constant, which turns insignificant.

In the column (3), we are interested in controlling for the stock market behaviour of asset  $i$ . It is entirely plausible that users create new submissions to discuss  $i$ ’s large returns, large volatility, or large trading volume in week  $t$ . To control for these patterns, we introduce the average excess rate of return in week  $t$ ,  $R_{i,t}^e$ , the average standard deviation in the excess rate of return,  $\sigma_{i,t}$ , and the average

dollar volume of shares traded in billions of USD,  $Q_{i,t}$ , using data from The Center for Research on Security Prices (CRSP). We detail the computation of each variable in Appendix A.3. Each of these controls is again expressed in terms of deviations from weekly averages across assets. Indeed, all are significant in explaining the number of new authors, but the improvement in  $R^2$  is marginal. The surprising finding is that the relationship between new authors and average price returns appears negative. Our parameters of interest,  $c$  and  $r$ , remain largely unchanged.

**Identifying the Contagion Mechanism: Matching on Observables** According to Table 1, asset discussions indeed seem to have an endogenous component. However, can WSB lend us more insight into the transmission mechanism from one passionate investor the next?

In order to validate our earlier result, we measured the propensity for a user to post a submission about a ticker, conditional on them previously commenting on submissions discussing the same ticker. The main challenge comes in addressing the network’s endogenous link formation, as hidden characteristics will determine a user’s *choice* to comment. This problem, known as ‘homophily’ in the broader network literature, leads to an overestimated spillover (Aral et al. 2009). We address this issue by matching on observable user characteristics. We filter the sample of submissions to the top one percent of tickers (top percentile tickers), by number of mentions, for this portion of our study. Individuals who commented on a top percentile ticker submission form the *treatment* group, and are matched to users who did not comment, the *control* group.

One important form of homophily to control for is individuals’ exposure to movements in the asset price: one may take interest in a stock because it experiences outsized returns, rather than hearing about it from someone else. To tackle this, we matched individuals who commented on a submission with those who did not comment, but were active on the forum *shortly after* the submission was made. This effectively controls for exposure to the same market environment, as well as associated news. In order to control for overall behavioural patterns, we matched individuals based on similar posting and commenting patterns, and on similar activity times on WSB. Lastly, to ensure the most accurate match, we consider individual’s interests in other subreddits, and match with those who share similar interests. Appendix A.4 offers further details on the distance measures used to compare users.

Following complex contagion theory, the spread of interests and beliefs may require multiple sources of activation, or social reinforcement (Lehmann & Ahn 2018). In order to quantify the level of social reinforcement necessary to drive a transmission in interest, users were grouped for every ticker by their commenting frequency on submissions related to that ticker. About 90% of users commented on fewer than five different submissions about a specific ticker; we, therefore, grouped together those who commented on five or more submissions. We matched each group separately to different control groups. Regardless of the number of submissions they commented on, we do not include commenters on the same ticker in any control groups for that ticker. Henceforth, we consider the number of distinct submissions that an individual commented on, prior to their submission, the amount of treatment this individual received. This effectively yields five treatment groups, for people who commented on one, two, three, four, or over five different submissions. These categories are plotted on the bottom of Figure 6.

The result of interest is the difference in the observed proportion of submissions in the treatment and control groups,  $\pi_j^T - \pi_j^C$ , where

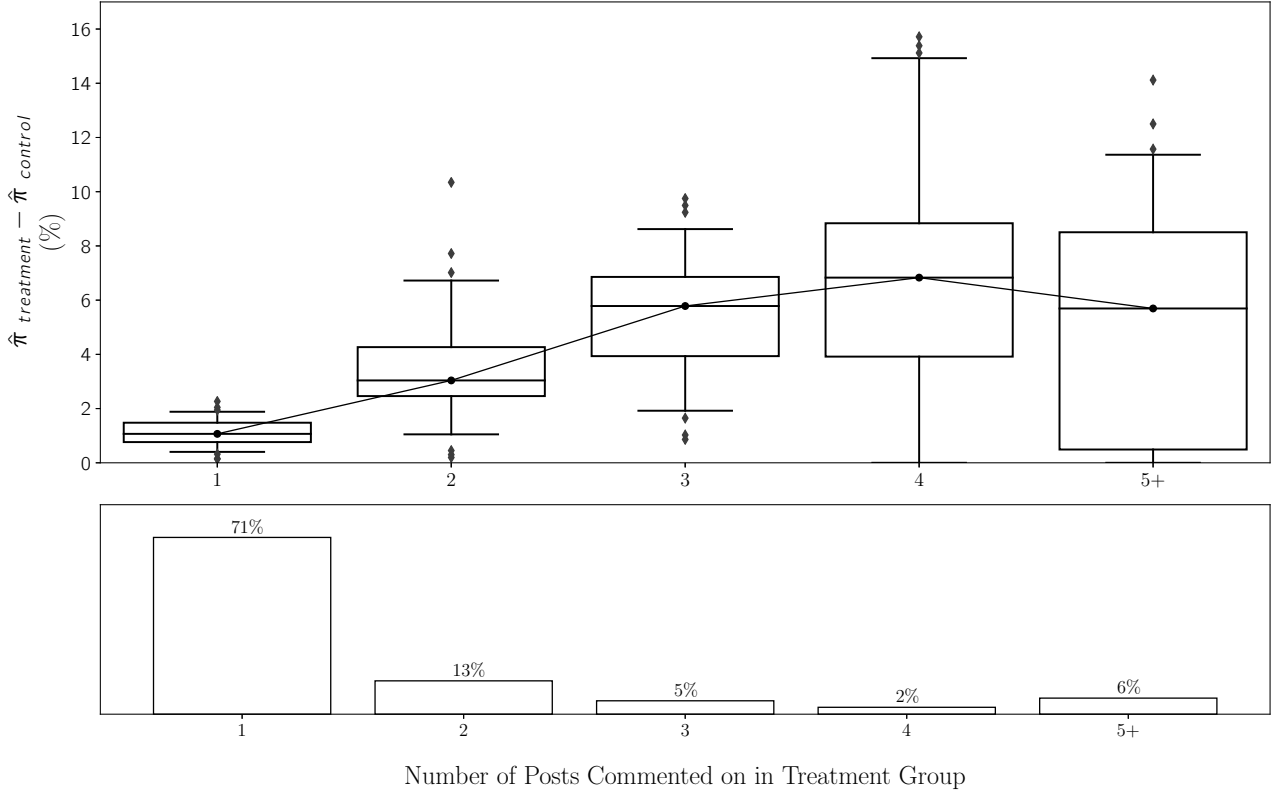


Figure 6: **Estimated Effect of Commenting on Posting a New Submission;** the distribution of the estimated treatment effect ( $\hat{\pi}_j^T - \hat{\pi}_j^C$ ) across the top one percent of tickers is graphed as a series of box plots. The solid line connects the median estimate, and the protruding dashes represent the 5<sup>th</sup> and 95<sup>th</sup> percentiles. The histogram below shows the proportion of individuals who comment between one and over five times on a submission for a given ticker before either creating a post themselves, or losing interest.

$$\pi_j^T = \mathbf{E}(\pi_{i,j,t+T} = 1 | l_{i,j,t} = 1),$$

$$\pi_j^C = \mathbf{E}(\pi_{i,j,t+T} = 1 | l_{i,j,t} = 0),$$

and  $\pi_{i,j,t+T}$  is one if user  $i$  posts a novel submission on asset  $j$  at some time  $t + T$ , and  $l_{i,j,t}$  is one if user  $i$  previously commented on a submission about asset  $j$  at time  $t$ .  $l_{i,j,t}$  is only observed once for each individual  $i$ . Therefore, we constructed the counterfactual,  $\hat{\pi}_j^C$ , from users matched to the commenters based on observable characteristics. The difference  $\hat{\pi}_j^T - \hat{\pi}_j^C$  thus measures the likelihood a user will post a submission given they are exposed to the asset on WSB, and not because of unobserved characteristics.

Figure 6 presents the observed differences in proportions as a function of the frequency with which a user commented. The first important result is that all differences in proportion are positive, implying that engaging in discussions on WSB about an asset strictly increases future interest in the asset. This effect becomes more prominent the more a user comments on submissions related to the same asset, suggesting some form of threshold contagion in asset demand. However, the small samples for users who commented four or over five times introduce large confidence boundaries.

The effect in Figure 6 appears small in absolute terms, however, we must consider that there is an order of magnitude more comments than submissions, as seen in Figure 2. On average, we observed a fourfold increase in the probability of authoring a new submission on an asset when an individual is exposed to a single submission on the same asset. The probability increases with the number of comments on distinct submissions. Further details on baseline probabilities for people who post in both the control and treatment groups are available in Appendix A.4.

### 3.2 Consensus Formation among Investors

Net of interest, the question still remains whether the WSB community actively forms a consensus on the direction an asset will take in the future. This section studies the network of user interactions on WSB in relation to the sentiments detected in user submissions. We find evidence that individual opinion formation is based, in part, on the opinions of peers, net of any price trajectory. This is indicative of consensus formation among investors.

**User Interaction Network** We formed user interaction networks in the following steps. First, we extracted all submissions that contain a single mention of an asset; these count as nodes. We placed a directed edge, from an earlier submission to a later one if the author of the later submission commented on the earlier submission. Submissions with the same author are not linked - an author's own previous submissions about a ticker are considered as a separate, independent variable when evaluating network effects. The resulting graph characterises the dynamic network through which active buying, or selling, decisions spread. If one individual comments on a submission about a ticker, prior to expressing their own view on this ticker, their revealed sentiment may be influenced by the initial submission they commented on. This section uses this network to track the flow of buying, or selling, preferences from one submission to the next.

**Individual Opinion Formation** The extent to which buying and selling decisions between linked submissions correlate, net of their relationship to actual price movement in the underlying asset, can be used to approximate a spillover from social interaction. We define  $\phi_i$  as *i's opinion*, or *sentiment*, which embodies their expectation of the *future* price trajectory of an asset, reminiscent of the extrapolation exercise that ordinary investors perform in the model of Barberis et al. (2018).

The following multinomial logit problem isolates the spillover coefficients of interest using a standard spatial lag model:

$$\log\left(\frac{P(\phi_{i,t} = s)}{P(\phi_{i,t} = 0)}\right) = \underbrace{m + aR_t^e - a'\widetilde{R}_t^e - b\sigma_t + \alpha\phi_{i,t-T}}_{\text{baseline model}} + \underbrace{\sum_s \beta^s \sum_{j \neq i}^N w_{i,j} \phi_{j,t-T}^s}_{\text{spatial lag}} + \varepsilon_{i,t}, \quad (2)$$

where the dependent variable is the log probability of recording sentiment  $s \in \{+, -\}$  at time  $t$  for author  $i$ , over the neutral benchmark  $s = 0$ .  $R_t^e$  is the daily price return in excess of the market return of the asset being discussed;  $\widetilde{R}_t^e$  is the weekly excess return;  $\sigma_t$  is the standard deviation in excess returns;  $m$  is a constant term. We measure  $\sigma_t$  as the standard deviation in excess returns,  $R_t^e$ , in the ten days preceding the timestamp of  $i$ 's submission (See Appendix A.3 for details). We drop a subscript denoting that the submission and all related variables are calculated for a specific ticker since each network refers to a specific asset; this is also done for Eq. 3 later in this section.

The persistence of  $i$ 's sentiment is separated in the form of a lag term  $\alpha\phi_{i,-1}$ , where the subscript  $-1$  indicates a previous time period.  $\phi_{i,-1}$  is a set of dummies:  $\phi_{i,-1}^{NA} = 1$  if the author has not posted before,  $\phi_{i,-1}^s = 1$  if sentiment  $s$  is expressed in an author's previous submission about the asset. These covariates, in isolation, form the baseline model for user sentiment, as indicated in Eq. 2. Further details on the timing of  $t$ , are covered in Appendix A.3.

The spatial lag model augments this baseline model by introducing  $W$ , the normalised adjacency matrix for the network, described at the start of this section. Elements in  $W$ ,  $w_{i,j}$ , equal the number of comments made by author  $i$  on an older submission by a different author  $j$  with the same extracted ticker, divided by the total number of comments made by  $i$ 's author on all older submissions with the same extracted ticker. With neutral sentiments as the benchmark case, scalar  $\beta^s$  is the log-odds of a submission expressing sentiment  $\phi_i = s$ , given that fraction  $w_{i,j} \leq 1$  neighbours also express sentiment  $s$ . For presentation, we use matrix notation  $W\phi^s$  instead of the summation term in Eq. 2. Finally, persistent average sentiments are further captured by the constant  $m$ , and  $\varepsilon_{i,t}$  denotes the error term.

Table 2: WSB Sentiment Spillovers

|                    | <i>Dependent variable:</i> |                      |                      |                      |                      |                      |
|--------------------|----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                    | $\phi_i = -1$<br>(1)       | $\phi_i = +1$<br>(2) | $\phi_i = -1$<br>(3) | $\phi_i = +1$<br>(4) | $\phi_i = -1$<br>(5) | $\phi_i = +1$<br>(6) |
| $R^e$              | -.46*** (.07)              | .03 (.03)            | -.46*** (.07)        | .02 (.03)            | -.47*** (.07)        | .02 (.03)            |
| $\widetilde{R}^e$  | -.21*** (.04)              | .04*** (.02)         | -.20*** (.04)        | .04*** (.02)         | -.19*** (.04)        | .03** (.02)          |
| $\sigma$           | -.84*** (.16)              | -.95*** (.10)        | -.86*** (.16)        | -.97*** (.10)        | -.96*** (.16)        | -1.00*** (.11)       |
| $\phi_{-1}^+$      | -.44*** (.03)              | .45*** (.02)         | -.42*** (.03)        | .46*** (.02)         | -.43*** (.04)        | .48*** (.02)         |
| $\phi_{-1}^0$      | -.80*** (.03)              | -.38*** (.02)        | -.78*** (.03)        | -.37*** (.02)        | -.81*** (.03)        | -.39*** (.02)        |
| $\phi_{-1}^-$      | .72*** (.04)               | -.17*** (.04)        | .72*** (.04)         | -.16*** (.04)        | .76*** (.04)         | -.16*** (.04)        |
| $\phi_{-1}^{NA}$   | -.22*** (.02)              | .04*** (.01)         | -.23*** (.02)        | .02 (.01)            | -.23*** (.02)        | .03* (.01)           |
| $W\phi^+$          |                            |                      | -.12*** (.04)        | .09*** (.03)         | -.16*** (.04)        | .08*** (.03)         |
| $W\phi^0$          |                            |                      | -.21*** (.04)        | -.19*** (.03)        | -.25*** (.04)        | -.19*** (.03)        |
| $W\phi^-$          |                            |                      | .39*** (.05)         | -.20*** (.05)        | .35*** (.06)         | -.19*** (.05)        |
| $WR^e$             |                            |                      |                      |                      | -.16 (.20)           | .20** (.09)          |
| $W\widetilde{R}^e$ |                            |                      |                      |                      | -.06 (.06)           | .02 (.03)            |
| $W\sigma$          |                            |                      |                      |                      | .93* (.48)           | .04 (.36)            |
| $m$                | -.73*** (.01)              | -.07*** (.01)        | -.71*** (.02)        | -.04*** (.01)        | -.71*** (.02)        | -.04*** (.01)        |
| Model              | Baseline                   |                      | Spatial Lag          |                      | Spatial Durbin       |                      |
| AIC                | 193,749.7                  |                      | 193,573.9            |                      | 188,355.5            |                      |
| McFadden's $R^2$   | .01                        |                      | .01                  |                      | .01                  |                      |
| N                  | 95,981                     |                      | 95,981               |                      | 93,412               |                      |

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Notes: This table presents estimated log-odds coefficients for three multinomial logit models on the expressed sentiment in WSB submissions. Odd columns estimate the log-odds ratio of user  $i$ 's sentiment being bearish,  $\phi_i = -1$ . Similarly, even columns estimate the ratio of user  $i$ 's sentiment being bullish,  $\phi_i = +1$ . Neutral submissions serve as the benchmark.

An alternative model of interest is the spatial Durbin model, which includes spatial lags of co-variates in Eq. 2. As far as sentiment transmission goes, this model is useful because it considers the price movements of the discussed asset when the current author commented on the linked submission(s). The extra term's effect is measured by vector  $\delta$  in an augmented form of Eq. 2:

$$\log\left(\frac{P(\phi_{i,t}=s)}{P(\phi_{i,t}=0)}\right) = m + aR_t^e - a'\widetilde{R}_t^e - b\sigma_t + \alpha\phi_{i,t-T} + \sum_{j \neq i}^s \beta^s \sum_{j \neq i}^N w_{i,j} \phi_{j,t-T}^s + \underbrace{\delta \sum_{j \neq i}^N w_{i,j} \begin{bmatrix} R_{t-T}^e \\ \widetilde{R}_{t-T}^e \\ \sigma_{t-T} \end{bmatrix}}_{\text{spatial Durbin}} + \varepsilon_{i,t}. \quad (3)$$

Table 2 presents the results of the multinomial models specified in Eqs. 2-3. For further details, we outline the construction of each variable in Appendix A.3. The baseline model in columns (1) and (2) suggests that submissions are more likely to be bearish versus neutral on days where the excess return of the asset is low. The estimated coefficient of -0.46 translates to an increase in the probability of the submission being bearish by almost 5% if the asset price drops by 10% in excess of the market return on the day that the submission is made. The converse for bullish posts is smaller and statistically insignificant. The return over a trading week is a significant variable; a -10% accumulated return over the preceding five trading days increases the probability of a submission being bearish by 2.1%, and reduces the probability of a submission being bullish by 0.4%. Moreover, the standard deviation in daily returns observed in the previous five trading days strongly increases the likelihood that a submission expresses neutral opinions. Overall, sentiments follow current and recent changes in the asset prices, but high volatility increases users' expressed uncertainty on the future trajectory of the asset.

Beyond price performance, authors' sentiments persist over time. Submissions are 56.8% more likely to express bullishness if the same author previously made a bullish post about the ticker, all else equal. The equivalent effect for bearishness is more significant; a submission is more than twice as likely to be bearish if its author posted a bearish submission in the past. However, authors who post a submission for the first time, or expressed equivocal sentiments in their previous post, are significantly more likely to express a neutral opinion.

These estimates for the baseline do not vary substantially across the various specifications. We present the remaining coefficients estimated for Eq. 2 in columns (3) and (4). Departing from the baseline model, the data demonstrates that authors who previously commented on a bearish post are 47.7% more likely to express bearish over neutral sentiments, and 18.1% less likely to express bullish sentiments over neutral sentiments. Similarly, but less markedly, authors who previously commented on at least one bullish submission are 9.4% more likely to write a bullish submission, yet 11.3% less likely to write a bearish one. Comparable results are also observed for neutral posts.

The results offer strong evidence of the endogenous spread of sentiments between users. The effect is more pronounced for assets with bearish outlooks. One hypothesis is that users rely on discussions to time a potentially lucrative downturn in the price of an asset. Another is that panic spreads faster, given that volatility is a big driver of uncertainty expressed in submissions. This is, in part, supported by the spatial Durbin model detailed in Eq. 3. The estimated coefficients in columns (5) and (6) imply that a submission is more likely to be bearish if the author commented on a submission made on a day the asset's price was volatile. However, the coefficients in this specification are not statistically significant. The precise driver of the asymmetry in buying versus selling decisions requires deeper exploration.

**Aggregate Opinion Formation** One pertinent question is on the persistence of aggregate opinions. This is of particular importance when considering market impact: do investors come to an agreement



Table 3: WSB Sentiment Spillovers from the Leontieff Inverse

|  | <i>Dependent variable:</i> |                      |                      |                      |
|--|----------------------------|----------------------|----------------------|----------------------|
|  | $\phi_i = -1$<br>(1)       | $\phi_i = +1$<br>(2) | $\phi_i = -1$<br>(3) | $\phi_i = +1$<br>(4) |
| $(I - W)^{-1}\phi^+$                       | -.13*** (.04)              | .11*** (.03)         | -.08 (.07)           | .08* (.05)           |
| $(I - W)^{-1}\phi^0$                       | -.22*** (.04)              | -.13*** (.03)        | -.23*** (.08)        | -.14*** (.05)        |
| $(I - W)^{-1}\phi^-$                       | .30*** (.06)               | -.10* (.05)          | -.001 (.11)          | -.24*** (.09)        |
| $(I - W)^{-1}R^e$                          | -.07 (.17)                 | .10 (.08)            | -.06 (.17)           | .10 (.08)            |
| $(I - W)^{-1}\widetilde{R}^e$              | -.05 (.04)                 | -.01 (.02)           | -.05 (.04)           | -.01 (.02)           |
| $(I - W)^{-1}\sigma$                       | 1.24** (.51)               | .36 (.36)            | 1.19** (.51)         | .34 (.36)            |
| $(I - W)^{-1}\phi^+ \times \phi_{-1}^{NA}$ |                            |                      | -.08 (.08)           | .04 (.05)            |
| $(I - W)^{-1}\phi^0 \times \phi_{-1}^{NA}$ |                            |                      | .01 (.09)            | .01 (.06)            |
| $(I - W)^{-1}\phi^- \times \phi_{-1}^{NA}$ |                            |                      | .43*** (.13)         | .21** (.11)          |
| $m$  | -.70*** (.02)              | -.04*** (.02)        | -.69*** (.02)        | -.03* (.02)          |
| AIC  | 169,963.5                  |                      | 169,961.3            |                      |
| McFadden's $R^2$                           | .01                        |                      | .01                  |                      |
| N  | 84,949                     |                      | 84,949               |                      |

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

*Notes:* This table presents estimated log-odds coefficients for two multinomial logit models on the expressed sentiment in WSB submissions, using extra variables to model bullish/bearish sentiment formation in Table 2. Odd columns estimate the log-odds ratio of user  $i$ 's sentiment being bearish,  $\phi_i = -1$ . Similarly, even columns estimate the ratio of user  $i$ 's sentiment being bullish,  $\phi_i = +1$ . Neutral submissions serve as the benchmark. The coefficients for the baseline variables, which include the excess daily return, excess weekly return, volatility, and old sentiments by the same author, are not documented here, since they do not change meaningfully.

about whether an asset should be bought or sold? We estimate this phenomenon through a parameter  $\alpha$ , which denotes the aggregate buying/selling intensity resulting from individual opinion formation.

One way gauge  $\alpha$  is to replace matrix  $W$  in Eqs. 2-3 by its Leontief inverse,  $(I - W)^{-1}$ . Instead of a weighted average of  $i$ 's direct neighbours, the Leontief inverse returns the overall position of user  $i$  in the network. While immediate neighbours are still highly weighted, the sentiments of their neighbours, and those neighbours' neighbours, feature as well. The end result is that  $(I - W)^{-1}$  captures the sentiments related to an asset in the broader context of the whole forum, appropriately discounting those further removed from  $i$ . Therefore, the Leontief inverse is better suited in measuring long term fluctuations in aggregate sentiment.

The first two columns in Table 3 show the coefficients of Eq. 3, but with the adjacency matrix replaced by its Leontief inverse, and the diagonal elements set to zero. The baseline coefficients are suppressed, since they do not vary substantially from Table 2. The stark difference here is the much stronger effect of volatility,  $(I - W)^{-1}\sigma$ . It reflects a strong persistence of bearish sentiments in discussions when the asset in question experiences volatility. In contrast to the statistically insignificant coefficient on overall returns,  $(I - W)^{-1}R^e$ , volatility, and the panic it appears to induce, has a powerful, propagating effect in the formation of investor opinions on WSB.

The final set of columns in Table 3 investigates whether sentiments detected in users who post their first submission, versus those that have already expressed an opinion in the past, are noticeably

different. The answer is in the affirmative; interacting the relevant dummy variable with the spillover term,  $(\mathbf{I} - \mathbf{W})^{-1} \phi^s \times \phi_{-1}^{NA}$ , demonstrates that the large bearish spillovers take hold in users who post for the first time. The original coefficient on  $(\mathbf{I} - \mathbf{W})^{-1} \phi^-$  is statistically insignificant in predicting bearish over neutral sentiment.

## 4 Contagion, Consensus, and Asset Bubbles

Table 4: WSB Sentiments and Asset Price Movements

|                         | <i>Dependent variable:</i> |                 |                             |
|-------------------------|----------------------------|-----------------|-----------------------------|
|                         | $R_{i,t}^e$                | $\log(V_{i,t})$ | $\log(P_{i,t}^H/P_{i,t}^L)$ |
|                         | (1)                        | (2)             | (3)                         |
| $\Phi_{i,t}^+$          | 0.004 (0.0001)             | 0.15 (0.004)    | 0.01 (0.0001)               |
| $\Phi_{i,t}^0$          | 0.002 (0.0001)             | 0.09 (0.003)    | 0.01 (0.0001)               |
| $\Phi_{i,t}^-$          | -0.01 (0.0002)             | -0.13 (0.01)    | -0.01 (0.0002)              |
| Within-ticker           | Yes                        | Yes             | Yes                         |
| Within-day              | Yes                        | Yes             | Yes                         |
| N                       | 14.4m                      | 14.4m           | 14.4m                       |
| Adjusted R <sup>2</sup> | .0002                      | .0005           | .001                        |

*Notes:* This table presents OLS estimators for the correlation between aggregated daily sentiments,  $\Phi_{i,t}^s$ , on WSB, and three stock market indicators as dependent variables. All variables are ‘within’-transformed to denote their deviations from cross-sectional and temporal means. Subscript  $i$  denotes the asset, and  $t$  the trading day. Superscripts +, 0 and - denote bullish, neutral and bearish sentiments, respectively. Corresponding standard errors are in parentheses. Column (1) uses daily returns, calculated as the ratio of subsequent closing prices for assets  $i$  at times  $t$  and  $t - 1$ , as recorded in CRSP, minus the risk free rate, as reported in K. French’s data library. Columns (2) and (3) regress the variation in sentiment within asset  $i$  and time  $t$  on the within asset and time variations in log of trading volume. The same is done in column (3) for the log-difference in daily intra-day high and low prices for asset  $i$ .

Do social interactions cause asset price fluctuations? In Section 3 we find that individuals rely on their social connections for investment advice. Assets gain prominence on WSB over the course of months, driven by a self-reinforcing social mechanism within the forum itself, rather than a change in fundamentals. Herding and social dynamics appear to play a significant role in what people consider for their investment portfolio. Discussions on WSB also impact the perception of future asset returns. Users on the WSB forum change their sentiments about assets depending on the views expressed by their peers. There is evidence that individuals come together and, on aggregate, reach consensus, creating an environment where coordination and market impact is possible.

In this section, we begin by offering empirical evidence of the connection between sentiment on WSB and asset returns. We subsequently propose a model for stock price changes due to social contagion. We theoretically determine regimes which result in asset price stability and in volatility. Simulations of the proposed model allow us to identify how asset prices evolve over time. Finally, we extract a causal effect of WSB on financial markets through a Two-Stage-Least-Squares approach modeling the effect of posting activity on trade volumes.

**Sentiment on WSB and Asset Prices** As a simple exercise, we define the total sentiment  $s$  on asset  $i$  at time  $t$ ,  $\Phi_{i,t}^s$ , as the number of submissions expressing sentiment  $s$  on the future price of asset

*i.* To gauge the relationship between sentiments and the stock market, we are interested in asset prices, volumes and volatilities. We proxy these with i) asset  $i$ 's daily rate of return in excess of the risk-free rate,  $R_{i,t}^e$ , ii)  $i$ 's daily log trading volume,  $\log(V_{i,t})$ , and iii)  $i$ 's log intra-day high-to-low price range,  $\log(P_{i,t}^H/P_{i,t}^L)$ . Full details for the variable's data sources and construction are available in Appendix A.3. Since all variables display substantial cross-sectional and temporal heterogeneity, instead of considering the raw variables described above, we look at their *deviation* from their cross-sectional means at time  $t$ , as well as the temporal means for each asset  $i$ , as displayed in Table 4, where the cross-sectional and temporal means, denoted by a bar, are subtracted from each variable. This method is identical to introducing fixed effects.

Table 4 displays the results of a simple linear regression model, using least-squares fits, for each of our stock market variables as dependent variables. Column (1) shows that sentiments  $\Phi_{i,t}^s$  correlate significantly with asset  $i$ 's daily excess returns: bullish posts relate to a positive return, neutral posts to a somewhat lower positive return, and bearish posts to a negative return. The coefficient for bearish returns is more pronounced: ten new bearish submissions on one asset, relative to the cross-sectional and historical averages, relate to a 10% downturn in its price. Columns (2) and (3) present two additional results of interest. Aggregate sentiments correlate significantly with the associated asset's daily trading volume, as well as the intra-day trading range. Users on WSB display a propensity to discuss volatile assets, with prospects of high returns, yet high risk. Interestingly, the number of bearish posts correlates with lower relative trading volumes and intra-day price ranges, suggesting some preference for bullish strategies.

These results do not indicate causation. While submissions likely raise attention to these assets, alone they fail to gauge whether these movements follow fundamental shifts in the underlying value, or simply an idiosyncratic rise in the demand for that asset.

## 4.1 Basic Framework

Table 4 demonstrates the strong connection between investor discussions on WSB and stock price movements, but can these discussions shed light on the drivers behind asset price fluctuations? Behavioural economists have long deliberated the role that social interactions play in investment decisions. This section models the key dynamics which emerge from social contagion to structure how assets respond to social dynamics.

The best context to place the behaviour observed on WSB is within the behavioural finance framework, separating market participants into 'smart' and 'ordinary' investors, as proposed in Shiller (1984). The field largely evolved to understand the economic decision making process of ordinary investors. Examples include Barberis & Huang (2009), who introduce a 'narrow framing' function to reproduce gambling tendencies often observed in the context of stock market bubbles. A more recent branch considers trend-following by ordinary investors, in what Barberis et al. (2018) label as 'extrapolation.' Hommes (2021) provides a thorough review of the many dimensions in which heterogeneous agents in stock markets are modelled, and the respective outcomes. The operating framework of such models is to investigate the interaction between types of agents in a market, typically some smart or ordinary investors, obeying straightforward rules. These models are somewhat distinct from a separate branch that seeks to understand 'social herding,' as discussed by Shiller (2005, 2017), which more closely relate to the behaviour on the WSB forum. Investor crowding and coordination is decomposed into two distinct effects in Kirman (1993). The interactions of groups of investors with different sentiments in the marketplace is discussed in Bouchaud & Potters (2003),

albeit at much shorter time scales than what is explored within this paper.

**Proposed Approach** Our goal is to evaluate how interactions affect asset price fluctuations, leveraging both our insights from WSB and the rich economic literature outlined above.

The intention is to incorporate the two key behaviours we observe empirically into a model of asset price changes: i) *contagion*, i.e. a size effect by which investors, motivated through social contagion, enter the market for one asset, and ii) *consensus*, i.e. an intensity effect capturing how well these investors coordinate their buying/selling decisions. We observe that WSB discussions are focused around single assets, rather than portfolios of assets: in submissions that contain a ticker, 77% only mention a *single* ticker. Because of this, as well as simplicity, we model a unique asset in isolation, and purposefully ignore broader market trends.

We are interested in separating out the behaviour of so-called ‘hype’ investors. Hype investors are short-sighted, enticed by the idea of a quick buck; they are susceptible to social contagion, driven, in part, by ‘gambler’s excitement’ *à la* Shiller (2005). The market impact of social contagion depends on how many investors could potentially be swayed by their peers ‘to buy into the hype’. In our model, we consider that a subset of investors is not swayed by social dynamics; they are grouped as ‘other’ investors.

The total market cap of an asset,  $pQ$ , is its price,  $p$ , times the number of shares outstanding,  $Q$ . It is equal to the sum of shares held by hype investors,  $Y$ , and other investors,  $S$ :

$$pQ = p(Y + S). \quad (4)$$

We attempt to model asset price changes in response to social contagion in the short term and, therefore, consider that the subset of non-hype investors choose to keep the nominal value invested in the given asset fixed:

$$\frac{d}{dt}[pS] = 0.$$

There is a second, implicit assumption that the number of shares,  $Q$ , does not change with time. For completeness, Appendix A.5 considers an alternative assumption whereby ‘other’ investors hold constant the number of shares held.

Taking the derivative of Eq. 4 with respect to time yields

$$\frac{\dot{p}}{p} = \frac{1}{pQ} \left( \frac{d}{dt} [pY(p)] \right), \quad (5)$$

where a dot denotes the derivative with respect to time  $t$  (we assume asset prices are differentiable). The left hand side of Eq. 5 is the instantaneous rate of change of the asset price, determined by hype investors’ asset demand on the right hand side.  $Y(p)$  is asset demand expressed as a function of price  $p$ . For a given shock in demand, Eq. 5 includes a ‘market depth’ adjustment  $1/pQ$ . Shocks to asset price, driven by an exogenous rise in asset demand from hype investors, are the subject of this paper.

Bouchaud & Potters (2003) argue that the impact of an idiosyncratic rise in asset demand  $Y(p)$  is linear on small time scales, where other investors fail to respond. Sustained for long enough, non-hype investors may also be prompted to buy into an asset experiencing a pronounced demand shock. We discuss developments on longer time scales in terms of demand elasticities among investors in Appendix A.5. These yield interesting patterns, whereby investors feed off of each others’ price responses. However, in this section we focus on the short-term developments as social contagion spreads among hype investors, and thus abstract away the interaction between different investor types.

## 4.2 Model with Contagion and Consensus

We incorporate the insights from Section 3 into a model capturing how the observed *consensus* and *contagion* dynamics impact asset prices. We consider that these dynamics govern the behaviour of hype investors,  $Y$ .

**Contagion** We turn to an adapted Kermack-McKendrick SIR epidemic model to quantify the evolution of asset demand via social contagion. This choice is justified through our results in Section 3.1 and empirically validated in Appendix A.7. We propose the following relationships between susceptible,  $V$ , active/infected,  $A$ , and recovered,  $B$ , investors in a specific asset:

$$\begin{aligned}\frac{dV}{dt} &= -cAV, \\ \frac{dA}{dt} &= cAV - rA, \\ \frac{dB}{dt} &= rA,\end{aligned}\tag{6}$$

where  $c$  is the contagion rate, and  $r$  is the recovery rate. The total population of potential hype investors,  $N = V + A + B$ , is assumed constant. We take this model literally to understand contagion in asset interest on WSB. ‘Active’ users forgo passive investment, adopt a specific position in an asset, and convince others to do the same. ‘Recovered’ users lose interest and cease to believe in the hype, and thus close their asset-specific position. ‘Susceptible’ users have not yet taken any investment action but could potentially become interested in the asset, as they seek investment advice from their peers. We use  $A$  as an estimate for the number of hype investors in an asset at a certain point in time.

**Consensus** The contagion model does not tackle the question of how active, hype investors form a consensus on their expectation of the future price trajectory. In other words, it counts the number of hype investors influenced by social contagion, but does not elaborate on their individual strategies.

To understand the buy/sell decisions influenced by social interaction, we turn to a model inspired by Bouchaud & Potters (2003). Consistently with Section 3.2, we define  $\phi_i$  as  $i$ ’s opinion, which embodies their expectation of the future price trajectory. We propose that individual investors use a non-linear decision rule with respect to an individual’s ‘information signal’  $\chi_i$  to form a buy or sell opinion about an asset:

$$\phi_i = f(\chi_i) = f\left(m_i + (a - a')\frac{\dot{p}}{p} - b\left(\frac{\dot{p}}{p}\right)^2\right),\tag{7}$$

where

- $\phi_i$  is the direction (buy/sell) that an investor chooses,
- $m_i$  is a term which captures social influence (its expanded form is given below),
- $a$  is the degree of trend following,
- $a'$  is a stabilising component,
- $b$  is an aversion for volatility.

The function  $f$  is a quantal response function:

$$f(\chi_i) = \begin{cases} +1, & \text{with probability } e^{\hat{\lambda}\chi_i}/(e^{\hat{\lambda}\chi_i} + 1) \\ -1, & \text{with probability } 1/(e^{\hat{\lambda}\chi_i} + 1), \end{cases}\tag{8}$$

with  $\chi_i$  defined as before,  $\hat{\lambda}$  specifying the amount of noise (or ‘irrationality’) in the decision process (with further discussions on  $\hat{\lambda}$  available in Bouchaud 2013).

Information unrelated to asset price, but related to the social dynamics, is modelled as

$$m_i = \beta \sum_{j=1}^N w_{i,j} \phi_j + \eta_i, \quad (9)$$

so that  $m_i$  is updated from a prior state  $\phi_i$  with weight  $w_{i,i}$ , combined with new information from neighbours  $j$ , weighted by  $w_{i,j}$ , and idiosyncratic news  $\eta_i$ . Elements  $w_{i,j}$  belong to the adjacency matrix  $\mathbf{W}$ , weighted by row totals, to produce a weighted average of information between social connections of  $i$  (consistent with the definition of  $\mathbf{W}$  in Section 3.2).  $w_{i,j}$  is greater than zero if a social connection exists between authors  $i$  and  $j$ ; it is zero otherwise. In the data, we estimate  $\mathbf{W}$  by observing whether the author of submission  $i$  commented on an older submission by  $j$ , for the same asset.

Our approach is justified empirically in Section 3.2. We find that the variables outlined above indeed have a statistically significant effect on the sentiment expressed by an individual about future asset performance.

**Aggregate Buying Intensity** Given a definition for  $\phi_i$ , we can derive the aggregate buying intensity  $\phi$  between all hype investors that are active in an asset at a given point in time. Since we are interested in the overall impact that hype investors have on price, deriving a cumulative consensus dynamic is of particular interest.

The formulation of individual  $\phi_i$  as in Eq. 8 yields the following functional form for  $\phi$ , computed as the average buy/sell decision across all active investors:

$$\phi = \frac{e^{\hat{\lambda}\hat{\chi}} - 1}{e^{\hat{\lambda}\hat{\chi}} + 1} = \tanh\left(\frac{\hat{\lambda}\hat{\chi}}{2}\right), \quad (10)$$

where the first equality stems from the fact that we are taking an average across identical Bernoulli random variables;  $\hat{\chi}$  is an estimate of the population average.

We substitute  $(\hat{a} - \hat{a}')$  and  $\hat{b}$  as population equivalents of the weights in Eq. 7. Individual buy/sell decisions are derived from the sentiments expressed by neighbors and aggregated by a single individual through the adjacency matrix  $\mathbf{W}$ , capturing how much an individual weights the opinion of his neighbors, as described in the context of Eq. 9. Since all  $\phi_i$ ’s are aggregated, the adjacency matrix  $\hat{\beta}\mathbf{W}$  is subsumed by a multiplier  $\hat{a}$ . The decision to subsume the adjacency matrix by a multiplier  $\hat{a}$  would yield an equality under certain assumptions on the connectivity matrix  $\mathbf{W}$ , or an approximation in other cases. For example, if we assume a fully connected network, weights  $w_{ij}$  in Eq. 9 are constant, allowing the connectivity to be perfectly captured by  $\hat{a}$ .

The aggregate buying intensity can be computed as:

$$\phi = \tanh\left[\frac{\hat{\lambda}}{2}\left((\hat{a} - \hat{a}')\frac{\dot{p}}{p} - \hat{b}\left(\frac{\dot{p}}{p}\right)^2 + \hat{a}\phi\right)\right]. \quad (11)$$

**Asset Price Returns Driven by Contagion, and Consensus** Armed with a model for how hype investors enter the market for a specific asset (our contagion model) and for how they reach an aggregate buying/selling intensity, we can proceed to model the price dynamics as a function of both contagion and consensus.

We re-write asset demand by hype investors as a function of time and price:

$$Y(p) = \frac{M}{p} A(t) \phi(p, t).$$

where  $M$  is hype investors' average purchasing power, assumed constant. Via substitution into Eq. 5, we find

$$\frac{\dot{p}}{p} = \frac{M}{pQ} \times \frac{d}{dt} [A(t) \phi(p, t)]. \quad (12)$$

The growth in asset demand is the growth in buying intensity and of new investors joining existing traders. New investors join according to our proposed contagion model, which varies with time. Consensus is formed both on the basis of time and asset price.

Taking the derivatives, and rearranging Eq. 12, yields the following first-order differential equation:

$$\frac{\dot{p}}{p} = \frac{MA}{pQ} \times \frac{\phi(cV - r) + \dot{\phi}}{1 + (1-s)\varepsilon_\phi}, \quad (13)$$

where

$$\dot{\phi} = \frac{\partial \phi}{\partial t} = \frac{\hat{\lambda} \hat{a}}{2} \operatorname{sech}^2 \left[ \frac{\hat{\lambda}}{2} \left( (\hat{a} - \hat{a}') \frac{\dot{p}}{p} - \hat{b} \left( \frac{\dot{p}}{p} \right)^2 + \hat{a} \phi \right) \right] \quad (14)$$

is the partial rate of change in buying intensity with respect to time,

$$\varepsilon_\phi = -\frac{p}{\phi} \frac{\partial \phi}{\partial p} = \frac{p \hat{\lambda}}{2 \phi} \left( (\hat{a} - \hat{a}') - \hat{b} \frac{\dot{p}}{p} \right) \operatorname{sech}^2 \left[ \frac{\hat{\lambda}}{2} \left( (\hat{a} - \hat{a}') \frac{\dot{p}}{p} - \hat{b} \left( \frac{\dot{p}}{p} \right)^2 + \hat{a} \phi \right) \right] \quad (15)$$

is the price elasticity of the buying intensity, and  $s = S/Q$  is the fraction of shares owned by non-hype investors. The precise steps taken in the derivation are outlined in Appendix A.5. Our model assumes that  $A, V$  vary slowly as compared to price. It also considers that price is continuous and can be differentiated.

Eq.13 determines the change in price as a function of the rate at which new investors enter the market, spurred by social contagion, and the rate of change in buying intensity. This allows us to isolate the impact that social contagion has on price dynamics, while capturing the two fundamental aspects of social contagion observed in the data, namely *contagion* and *consensus*. Further breaking down Eq. 13 highlights the intuition behind each component:

$$\frac{\dot{p}}{p} = \underbrace{\frac{MA}{pQ}}_{(1) \text{ Capacity}} \times \underbrace{\frac{1}{1 + (1-s)\varepsilon_\phi}}_{(2) \text{ Response}} \times \underbrace{(\phi(cV - r) + \dot{\phi})}_{(3) \text{ Demand}}. \quad (16)$$

Component (1), which we refer to as 'capacity,' is a market depth term, measuring how much market power hype investors have as a share of the total market cap of the the stock. Component (2), which we term 'response,' captures the price elasticity of hype investors weighted by the fraction of shares owned by other investors  $s$ . Let us consider a simple exercise with components (1) and (2). As the shares held by other investors approaches one ( $s$  goes to one), the price elasticity of the buying intensity of hype investors ceases to matter. The response term approaches one at this limit, and the price change is simply the extra money hype investors pour into the asset, divided by the total nominal value of all shares. At the other extreme, as  $s$  approaches zero, all new shares are bought from existing hype investors: the response term is now a function of the inverted price elasticity.

Term (3) captures the interplay of two important forces: the rate at which new hype investors enter the market for the asset and the overall change in buying intensity among the hype investors. The first part of (3),  $\phi(cV - r)$ , quantifies how many active hype investors enter the market for the asset. At the point of entry, these investors exert a buying intensity of  $\phi$ . The second part of (3),  $\dot{\phi}$ , captures the change in buying intensity.

### 4.3 Chaotic Prices from Coordinated Strategies

**Impact of  $\dot{\phi}$ ,  $\varepsilon_\phi$  on Price Stability** Our original derivations offer some insight into the primary components that impact price, as explained after Eq. 16. The influence of certain variables appears trivial to evaluate: for example, as  $A$  or  $M$  in Eq. 16 increase, social contagion will have greater market impact. However, the relevance of  $\dot{\phi}$ , in Eq. 14, and  $\varepsilon_\phi$ , in Eq. 15, is less obvious.

We turn to Hommes (2013) for an understanding of price stability, or instability. We consider a scenario when hype investors begin to discuss a stock but no price impact has been felt (setting  $\dot{p}/p = 0$ ) and analyse the price dynamics that ensue. This system produces a stable steady state at  $\phi = 0$  if  $\hat{\lambda}\hat{a}/2 \leq 1$ , implying that investors are not connected enough for sentiments to amplify. In this scenario, a small value for  $\hat{\lambda}\hat{a}/2$  implies that consensus is difficult to reach. However, with  $\hat{\lambda}\hat{a}/2 > 1$ , we arrive at a ‘pitchfork’ bifurcation, as described in Hommes (2013). The buying intensity produces two stable steady states,  $\phi^+ > 0$ ,  $\phi^- < 0$ , and one unstable steady state at  $\phi^0 = 0$ . These interesting dynamics illustrate the potentially important effect of strong feedback in  $\beta W$ , as investors coordinate to buy or sell the asset together.

**Discrete Time Simulation** We perform a discrete time simulation of our final model, presented in Eq. 13. The key pattern of interest is the effect of a transitory increase in demand driven by hype on the price of an asset. To this end, we simulate the following dynamical system:

$$\begin{aligned} A_{t+1} &= \max[A_t + cV_tA_t - rA_t, A_0], \\ V_{t+1} &= V_t - cV_tA_t, \\ B_{t+1} &= B_0 + (V_0 - V_{t+1}) + (A_0 - A_{t+1}) \\ \phi_{t+1} &= \frac{A_0}{A_{t+1}}\phi_0 + \frac{A_{t+1} - A_0}{A_{t+1}} \tanh\left[\frac{\lambda}{2}\left(a\left(\frac{p_t - p_{t-1}}{p_{t-1}}\right) - b\left(\frac{p_t - p_{t-1}}{p_{t-1}}\right)^2 + \alpha\phi_t\right)\right], \\ S_{t+1} &= Q - MA_{t+1}\phi_{t+1}, \\ p_{t+1} &= \frac{p_0S_0}{S_{t+1}}, \end{aligned}$$

where a subscript 0 denotes the starting value. The choices made to simulate the model include: i) some fixed amount of hype investors,  $A_0$ , are always present, with a buying intensity of  $\phi_0 = 0.1$ , and ii) non-hype investors fix the nominal value of their investment,  $p_0S_0$ , as discussed earlier in this section. In practice this implies that, hype investors enter the market to buy the asset (since  $\phi_0 > 0$ ), which other investors sell to them at a higher price. The price changes and coordination of buying strategies affect the intensity with which hype investors purchase the asset in subsequent periods. Gradually, they leave the market until only the starting amount  $A_0$  remain, and the price returns to its ‘true’ value.

Since the focus is on asset demand by hype investors, we simulate various parameters for coordination,  $\lambda$ , persistence,  $\alpha$ , price trend,  $a$ , and aversion to volatility,  $b$ . The remaining values were



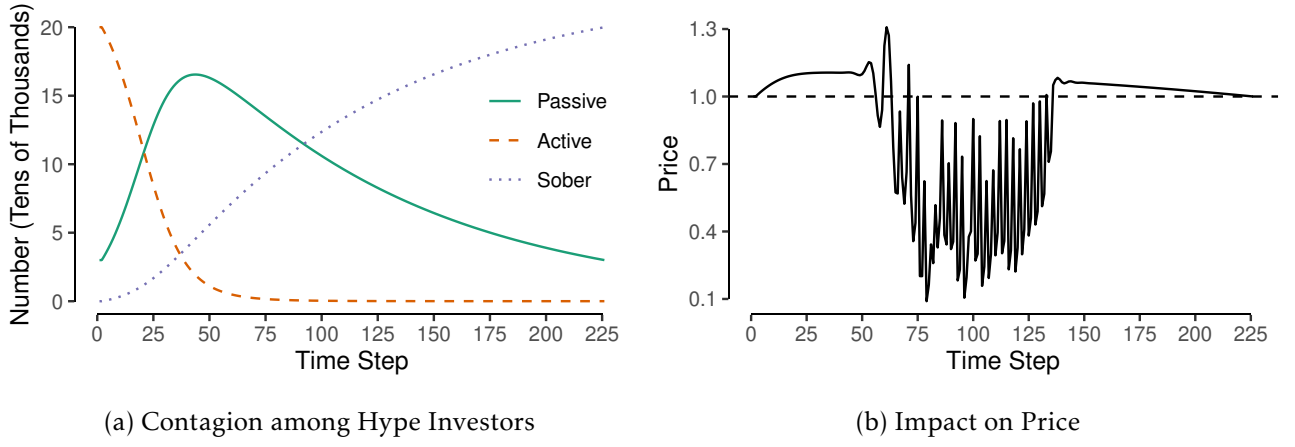


Figure 7: **Price Impact of Contagion and Opinion Dynamics**; on the left, the number of active investors in an asset increases steeply until time step 50, after which the number slowly declines to its initial value (as ‘active’ investors ‘recover’). On the right, the price first increases until time step 50, after which it experiences a volatile crash, then returns to its initial level. Parameter values are  $\lambda = 0.99$ ,  $\alpha = 0.45$ ,  $a = 0.2$ ,  $b = 0.5$ .

set to produce an equilibrium price of 1, and an equilibrium fraction of shares held by non-hype investors of 0.7: initial susceptible investors  $V_0 = 200,000$ , initial active investors  $A_0 = 30,000$ , initial recovered investors  $B_0 = 0$ , contagion rate  $c = 5 \times 10^{-7}$ , and recovery rate  $r = 0.01$ . The pattern of contagion produced is displayed in Figure 7a. At its peak, over 150,000 hype investors actively buy, or sell, the asset, before returning to the initial value after 225 time steps. For the price dynamics, the remaining inputs are for hype investor capital,  $M = 100$ , and number of shares,  $Q = 1,000,000$ .

Given the complex dynamics, qualitatively different patterns could be generated depending on the choice of demand parameters. For the baseline, we pick values that are close to the data:  $\lambda = 0.99$ ,  $\alpha = 0.45$ ,  $a = 0.2$ ,  $b = 0.5$ . The result, in Figure 7b, displays three important characteristics. The initial increase in demand, until time step 50, gradually drives up price. Once this growth is depleted, the price crashes dramatically, and extreme volatility persists until time step 150. At this stage, some active investors still remain in the market, but the price follows a steady downward trend until it reaches its initial level.

We test the prevalence of the slow initial price increase and sharp volatile drop across different initialization choices for  $\alpha$ ,  $\lambda$ ,  $a$ ,  $b$ . The results are presented in Appendix A.6. Certain elements of the price dynamics appear to change: the period of price volatility and magnitude (in percentage terms) of price drop. However, we observe that all initialization choices result in a slow, steady run-up in price followed by a crash and period of volatility.

**Key Implications for Price Stability** Our proposed model and simulations imply that social contagion has a destabilizing effect on markets. Excitement about an asset is slow to take off, however, coordination among investors gradually moves the price. This manifests in the slow run-up in price in our simulation, driven by the gradual increase in active hype investors due to contagion. As shown in Section 2.2, fear of downside spreads faster than optimism to the upside. Individuals eventually begin to doubt the value of an asset; panicked selling and volatility ensue.

Table 5: Relationship between WSB Activity and Financial Markets

|                         | <i>Dependent variable:</i> |                       |                      |                       |                   |                    |
|-------------------------|----------------------------|-----------------------|----------------------|-----------------------|-------------------|--------------------|
|                         | $R_{i,t}^e$                | $R_{i,t}^e$           | $\sigma_{i,t}$       | $\sigma_{i,t}$        | $Q_{i,t}$         | $Q_{i,t}$          |
|                         | (1)                        | (2)                   | (3)                  | (4)                   | (5)               | (6)                |
| $\Delta A_{i,t}$        | 0.0001*<br>(0.0000)        | 0.0001*<br>(0.0000)   | 0.001***<br>(0.0001) | 0.0005***<br>(0.0001) | 0.13***<br>(0.01) | 0.08***<br>(0.01)  |
| $\phi_{i,t-1}$          | -0.002**<br>(0.001)        | -0.002*<br>(0.001)    | 0.01***<br>(0.002)   | 0.01***<br>(0.001)    | 0.60***<br>(0.12) | 0.15***<br>(0.06)  |
| $\alpha$                | 0.001***<br>(0.0001)       | -0.001***<br>(0.0001) | 0.03***<br>(0.0001)  | -0.03***<br>(0.0002)  | 0.36***<br>(0.01) | -0.41***<br>(0.01) |
| Method                  | Pooled                     | Within                | Pooled               | Within                | Pooled            | Within             |
| R <sup>2</sup>          | 0.0004                     | 0.0004                | 0.01                 | 0.01                  | 0.22              | 0.29               |
| Adjusted R <sup>2</sup> | 0.0003                     | 0.0004                | 0.01                 | 0.01                  | 0.22              | 0.29               |

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

*Notes:* This table presents coefficients of the linear model in Eq. 17 for new authors on WSB discussing asset  $i$  in calendar week  $t$ ,  $\Delta A_{i,t}$ , and the average sentiment expressed about asset  $i$  in the previous week  $t - 1$ ,  $\phi_{i,t-1}$  fitted to data on three stock market variables: excess returns  $R_{i,t}^e$ , volatility in excess returns  $\sigma_{i,t}$ , and nominal trading volumes,  $v_{i,t}$ . All regressions use 64,698 observations. Models denoted as ‘Pooled’ under Method are estimated using a regular OLS. Where the Method is ‘Within’, an OLS regression is run on the within-week and within-ticker fluctuation in all dependent and independent variables.

#### 4.4 Evidence of Market Impact

Given our model in Section 4.3, we hypothesize that a causal relationship exists between the contagion mechanism within WSB and financial variables. In this section, we substantiate this hypothesis. We select three dependent market variables of interest: i) the average excess rate of return of asset  $i$  in week  $t$ ,  $R_{i,t}^e$ , ii) the standard deviation of excess returns in week  $t$ ,  $\sigma_{i,t}$ , and iii) the nominal trading volume of asset  $i$  in week  $t$ ,  $Q_{i,t}$ , in billions of USD. Our independent variables of interest are i) the number of new, unique authors posting about asset  $i$  in week  $t$ ,  $\Delta A_{i,t}$ , and ii) the average sentiment expressed about asset  $i$  in the previous week  $t - 1$ ,  $\phi_{i,t-1}$  (taking value  $-1$  if all are bearish, and  $+1$  if all are bullish). Our goal is to isolate how hype investor interest in an asset impacts market variables. For this section, we restrict the analysis to tickers mentioned over 30 times, and to submissions made from 2016 onward, for adequate coverage in our contagion variables.

We begin with a simple OLS approach to the following linear model:

$$Y_{i,t} = \alpha + \kappa \Delta A_{i,t} + \gamma \phi_{i,t-1} + u_{i,t}, \quad (17)$$

where  $Y_{i,t}$  is the target variable of choice,  $\alpha$  is a constant and  $u_{i,t}$  is the error term. The columns denoted by the ‘Pooled’ method in Table 5 report the estimated OLS coefficients from Eq. 17. In the columns labeled by the ‘Within’ method, all variables are de-measured with respect to time  $t$  and ticker  $i$ , to control for unobserved heterogeneity. Unlike in Table 1, all variables are also de-measured with respect to each ticker’s temporal average. This is identical to adding fixed effect dummies for

tickers and weeks.

The effects seem most pronounced for price volatility and nominal trading volume: ten new users are associated with an increase in the standard deviation of 0.005 in the within-group specification, and an average of \$80 million extra trading volume relative to other tickers or weeks. The coefficients on excess returns are small and not statistically significant, so there is little indication that WSB activity was able to push prices significantly in our sample.

Even though the OLS approach offers some intuition, it does not say much on the *causal effect* of WSB activity on the market, much like for Table 4. In Section 3.2, we show that sentiments expressed by individuals are formed both on the basis of social interactions, and recent stock returns and volatility. Therefore, there is a simultaneity problem: we are unable to gauge whether WSB activity affected movements in an asset at time  $t$ , or whether this WSB activity is driven by the very action in the asset we are trying to explain.

Table 6: 2SLS Estimate for the impact of WSB Activity on Assets

|                            | <i>Dependent variable:</i> |                    |                 |
|----------------------------|----------------------------|--------------------|-----------------|
|                            | $R_{i,t}^e$                | $\sigma_{i,t}$     | $Q_{i,t}$       |
|                            | (1)                        | (2)                | (3)             |
| $\Delta \widehat{A}_{i,t}$ | -0.0001** (0.0000)         | 0.0001*** (0.0000) | 0.09*** (0.01)  |
| $\phi_{i,t-1}$             | -0.002 (0.001)             | 0.01*** (0.001)    | 0.12** (0.05)   |
| $\alpha$                   | -0.001*** (0.0001)         | -0.03*** (0.0001)  | -0.41*** (0.01) |
| FS F-statistic             | 28.93                      | 28.93              | 28.928          |
| J-statistic                | 0.002                      | 88.702             | 0.115           |
| $R^2$                      | -0.001                     | 0.005              | 0.29            |
| Adjusted $R^2$             | -0.001                     | 0.005              | 0.29            |

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

*Notes:* This table presents coefficients of our 2SLS model in Eqs. 17-18 for the predicted new authors on WSB discussing asset  $i$  in calendar week  $t$ ,  $\Delta \widehat{A}_{i,t}$ , and the average sentiment expressed about asset  $i$  in the previous week  $t-1$ ,  $\phi_{i,t-1}$  fitted to data on three stock market variables: excess returns  $R_{i,t}^e$ , volatility in excess returns  $\sigma_{i,t}$ , and nominal trading volumes,  $v_{i,t}$ . All regressions use 64,698 observations. We use the number of authors who commented in the preceding three weeks, as well as average price returns in previous week, as instrumental variables for the number of new authors. The within-transformation is applied to all variables, with respect to tickers  $i$  and weeks  $t$ . FS F-statistic denotes the F-statistic on the first-stage regression, testing whether the coefficients on the instrumental variables are simultaneously equal to zero.

To answer the question of causality, we implement a rudimentary Two-Stage-Least-Squares (2SLS) estimator. This approach considers only the variation in WSB activity exogenous to *current* stock market movements as a predictive variable. Given the dynamics we observe in Section 3, past WSB activity is a strong predictor of current WSB activity.

The first-stage regression is

$$\Delta A_{i,t} = c + \beta_1 A_{i,t-1} + \beta_2 R_{i,t-1}^e + \varepsilon_{i,t}, \quad (18)$$

where  $\Delta A_{i,t}$  is the number of new authors who posted a submission discussing asset  $i$  in calendar week  $t$ ,  $A_{i,t-1}$  is the stock of unique authors who posted a submission discussing asset  $i$  in calendar

weeks  $t - 3$  to  $t - 1$ ,  $R_{i,t-1}^e$  is the average excess rate of return of asset  $i$  in calendar week  $t - 1$ ,  $c$  is a constant, and  $\varepsilon_{i,t}$  is an error term. We justify the choice of instruments as a reduced version of the specification in column (3) of Table 1. Specifically, there is the simple contagion effect stemming from the stock of users who discussed asset  $i$  in the preceding three weeks, and a sentiment effect captured by the change in price of asset  $i$  in the previous week.

The second stage is the target equation in Eq. 17, used to determine the impact of WSB on financial markets. The key detail is that,  $\Delta A_{i,t}$  takes the *fitted* values from the first-stage regression in Eq. 18, which we denote by  $\widehat{\Delta A}_{i,t}$ . In essence, these are the ‘hype’ investors expected to start conversations because of previous discussions they participated in, instead of any ‘news.’ The first key assumption is that the instruments in Eq. 18 are uncorrelated with the error term  $u_{i,t}$ . In other words, WSB activity in the past is not affected by market variables today. The second key assumption is that the stock of authors and lagged excess returns in Eq. 18 hold significant predictive power.

Table 6 presents the results from the 2SLS regression. The instrumental variables, as well as the dependent variables, are de-meanned by cross-sectional and temporal averages. The first-stage F-statistic, reported in the bottom part of Table 6 indicates that the instruments perform well in the 2SLS model.

The results presented in Table 6 are intuitive. As it stands, the little signal we get for price returns, in column (1), suggests a negative relationship between hype investors and price returns. While we try to control for sentiments in the previous week,  $\phi_{i,t-1}$ , the coefficient is statistically insignificant. One simple reason is that price returns are fast moving, whereas the contagion dynamics we model are slow. Discussions on any particular asset take months before peaking in size, as we demonstrate in Figure 5.

We would expect volatility to increase due to hype investors, which is supported by the results in column (2). However, judging by the J-statistic, our instruments do not appear valid in this particular instant. This is likely due to autocorrelation in asset price volatility, which is therefore codetermined with regards to past WSB activity.

Even though it is difficult to determine what direction a stock will move due to hype investors, one relationship is clear: as new investors enter the market for an asset, trading volumes increase. The relationship that is most pronounced from our 2SLS is that of new authors on WSB and nominal trading volume, as illustrated in column (3) in Table 6. This implies that our proposed contagion mechanism does, in fact, play a significant role in driving trading activity in an asset. As a reality check, we plot fitted values for two stocks, AMD and MU, in Figure 1, which highlights the trading activity capture by our model.

Furthermore, as dictated by our earlier model, the number of shares held by non-hype investors is an important factor which we do not account for in these regressions, except for a level-effect in the within-transformation. Simply, the impact  $\kappa$  in Eq. 17 may vary by  $i$ . Consider two simple examples. At the start of 2020, the WSB subreddit boasted approximately 800 thousand users<sup>11</sup>, giving a potential purchasing power of \$8 billion (assuming a purchasing power of \$10,000 for each WSB user<sup>12</sup>). Taking Apple Inc. as an example stock, even if all of WSB were to purchase this asset simultaneously, the price impact is unlikely to have been large, given Apple’s market cap of \$1.3 trillion, as of January 2<sup>nd</sup>, 2020. In contrast, GameStop had a market cap of approximately \$1.3 billion in January 2021, by which time WSB was frequented by 1.8 million users<sup>13</sup>, yielding a potential purchasing power of \$18 billion. Under a heavily simplified scenario, by which these users buy all shares available, this reflects an almost 14-fold increase in the price of GameStop equity (GME). At its highest price point, GME rose from \$19 to \$347 in January 2021, achieving an 18-fold

price increase.

Even though it is difficult to determine what direction a stock will move due to hype investors, one relationship is clear: trading volumes increase with new hype investors entering the market. The results in column (3) of Table 6 suggests that 100 new users in WSB, introduced via social contagion, lead to an average increase in trading volumes of \$9 billion, relative to other tickers and weeks. This implies that social contagion plays a significant role in driving trading activity in an asset. As an example, we plot fitted values for two stocks, AMD and MU, in Figure 1, highlighting that the model captures some important features of market activity of these two stocks. Of particular note are the cycles in hype investors, which fit the social contagion paradigm: excitement about AMD, for instance, rises over several weeks, then, as the number of users involved peaks, participants get bored and turn to discussing other assets. Eventually, a new catalyst, or the intake of new users, starts the cycle anew. Naturally, these users alone are unlikely to be the sole reason for elevated trading activities. Rather, WSB is likely a robust sample of retail investors' moods in a broader market setting, off of which institutional investors also conduct their own trades.

## 5 Conclusion

Our analysis documents two empirical findings: (i) 'contagion' and (ii) 'consensus' among retail investors. We test for contagion in asset interest in two different ways. Our results consistently show that investors become interested in discussing an asset, not because of fundamentals, but because other users discuss it. Subsequently, this paper tests whether an individual's sentiment about future asset performance are affected by those of others. We find that this is the case: people look to their peers to form an opinion about an asset's potential. Furthermore, this relationship is not symmetric: downside panic appears to spread from one individual to the next more quickly than enthusiasm about upside potential.

Is it possible to draw some conclusion about how asset prices change with respect to these dynamics? To that end, this paper introduces a model that decomposes investor asset demand due to social dynamics into i) hype around an asset caused by *contagion* and ii) opinion formation leading to overall *consensus* to buy/sell. The model indicates that, when investors are not very connected and consensus is difficult to reach, prices remains stable. However, as feedback among investors plays a greater role, the price impact is larger. This is verified through simulation, which indicates that the price will gradually increase as hype spreads, but eventually become highly unstable and crash. The price crash stems from panic selling, as investors turn nervous in the face of volatility. The connection between WSB and market activity is further validated using an instrumental variable approach. We find that WSB activity, exogenous to price movements, is able to explain significant variance in nominal trading volumes.

There are many questions that our analysis does not address. Causality is hard to establish, in the absence of investment records. A more detailed analysis, potentially directly linking discussions with investments by individuals is one area of future investigation. Recent research highlights that contagion and 'news shocks' result in distinct purchasing patterns on Amazon Deschates & Sornette (2005). It may be possible to decompose asset demand into 'hype-driven' and 'news-driven' demand offering a deeper understanding of what drives market volatility.

This research prompts us to inquire about the fate of 'homo economicus' – the rational individual who synthesizes all information. The observed behaviours within WSB can shed light on whether traditional economic assumptions are empirically valid. Considering that individuals look towards

their immediate neighbours to determine lucrative investment opportunities, with hype as an important factor in decision-making, the ‘homo economicus’ may lack some important attributes to describe the modern investor. Finally, the financial markets do not exist in isolation: investor decisions have broader implications on the cost of capital. It is important to evaluate how the influence of hype investors will percolate through the economy.

As social media galvanizes a larger pool of retail investors with the potential for exciting stock market gambles, it is crucial to understand how social dynamics can impact asset prices. With the first publicly acclaimed victory of Main Street over Wall Street, in the form of the GameStop short squeeze, it is unlikely that socially-driven asset volatility will simply disappear. In fact we observe the opposite: WSB grew from approximately 1.8 million users at the start of January 2021 to over 10 million users in April 2021.<sup>14</sup> Overall, the newly available data on WSB and changing investing climate, driven by increased social media usage, offer promising opportunities for future research.

## Notes

<sup>1</sup><https://datareportal.com/reports/digital-2021-global-overview-report>

<sup>2</sup><https://www.alexa.com/topsites>

<sup>3</sup><https://www.reddit.com/r/wallstreetbets/>

<sup>4</sup><https://www.nytimes.com/2021/02/04/business/media/new-york-times-earnings.html>

<sup>5</sup><https://www.alexa.com/topsites>

<sup>6</sup><https://redditblog.com/2019/12/04/reddits-2019-year-in-review/>

<sup>7</sup><https://subredditstats.com/r/wallstreetbets>

<sup>8</sup><https://andriymulyar.com/blog/how-a-subreddit-made-millions-from-covid19>

<sup>9</sup><https://www.reddit.com/r/wallstreetbets/>

<sup>10</sup><https://pushshift.io/>

<sup>11</sup><https://subredditstats.com/r/wallstreetbets>

<sup>12</sup><https://andriymulyar.com/blog/how-a-subreddit-made-millions-from-covid19>

<sup>13</sup><https://subredditstats.com/r/wallstreetbets>

<sup>14</sup><https://subredditstats.com/r/wallstreetbets>

## References

- Aral, S., Muchnik, L. & Sundararajan, A. (2009), ‘Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks’, *Proceedings of the National Academy of Sciences* **106**(51), 21544–21549.
- Barberis, N., Greenwood, R., Jin, L. & Shleifer, A. (2018), ‘Extrapolation and bubbles’, *Journal of Financial Economics* **129**(2), 203–227.
- Barberis, N. & Huang, M. (2009), ‘Preferences with frames: A new utility specification that allows for the framing of risks’, *Journal of Economic Dynamics and Control* **33**(8), 1555 – 1576.
- Bollen, J., Mao, H. & Zeng, X. (2011), ‘Twitter mood predicts the stock market’, *Journal of Computational Science* **2**(1), 1 – 8.
- Bouchaud, J.-P. (2013), ‘Crises and collective socio-economic phenomena: simple models and challenges’, *Journal of Statistical Physics* **151**(3), 567–606.
- Bouchaud, J.-P. & Potters, M. (2003), *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2 edn, Cambridge University Press.

- Christakis, N. A. & Fowler, J. H. (2008), 'The collective dynamics of smoking in a large social network', *New England journal of medicine* **358**(21), 2249–2258.
- Deschatres, F. & Sornette, D. (2005), 'Dynamics of book sales: Endogenous versus exogenous shocks in complex networks', *Physical Review E* **72**(1), 016112.
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018), 'Bert: Pre-training of deep bidirectional transformers for language understanding', *arXiv preprint arXiv:1810.04805*.
- Garcia, D. (2013), 'Sentiment during recessions', *The Journal of Finance* **68**(3), 1267–1300.
- Gentzkow, M., Kelly, B. & Taddy, M. (2019), 'Text as data', *Journal of Economic Literature* **57**(3), 535–74.
- Hommes, C. (2013), *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*, Cambridge University Press.
- Hommes, C. (2021), 'Behavioral and experimental macroeconomics and policy analysis: A complex systems approach', *Journal of Economic Literature* **59**(1), 149–219.
- Kirman, A. (1993), 'Ants, rationality, and recruitment', *The Quarterly Journal of Economics* **108**(1), 137–156.
- Lehmann, S. & Ahn, Y.-Y. (2018), *Complex spreading phenomena in social systems*, Springer.
- Leng, Y., Dong, X., Moro, E. & Pentland, A. S. (2018), *The Rippling Effect of Social Influence via Phone Communication Network*, Springer International Publishing, Cham, pp. 323–333.
- Ridenhour, B., Kowalik, J. M. & Shay, D. K. (2018), 'Unraveling r 0: Considerations for public health applications', *American journal of public health* **108**(S6), S445–S454.
- Shiller, R. J. (1984), Stock prices and social dynamics, Technical Report 2, The Brookings Institution.
- Shiller, R. J. (2005), *Irrational Exuberance: (Second Edition)*, Princeton University Press.
- Shiller, R. J. (2017), 'Narrative economics', *American Economic Review* **107**(4), 967–1004.
- Tetlock, P. C. (2007), 'Giving content to investor sentiment: The role of media in the stock market', *The Journal of finance* **62**(3), 1139–1168.
- Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D. & Nyhan, B. (2018), 'Social media, political polarization, and political disinformation: A review of the scientific literature', *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*.
- Yan, X., Guo, J., Lan, Y. & Cheng, X. (2013), A biterm topic model for short texts, in 'Proceedings of the 22nd International Conference on World Wide Web', WWW '13, Association for Computing Machinery, New York, NY, USA, p. 1445–1456.

## A Appendix

### A.1 Tickers Mentioned on WSB

Table 7: Most Frequent Ticker Mentions

| Ticker | Name                                   | Comment Count | Submission Count | Sum     |
|--------|--|---------------|------------------|---------|
| SPY    | S&P 500 Index                          | 243,572       | 7,925            | 251,497 |
| AMD    | Advanced Micro Devices, Inc.           | 121,249       | 5,555            | 126,804 |
| TSLA   | Tesla, Inc.                            | 107,987       | 5,419            | 113,406 |
| MU     | Micron Technology, Inc.                | 84,301        | 3,879            | 88,180  |
| AAPL   | Apple Inc.                             | 44,588        | 1,794            | 46,382  |
| AMZN   | Amazon.com, Inc.                       | 41,316        | 1,441            | 42,757  |
| SNAP   | Snap Inc.                              | 39,509        | 1,988            | 41,497  |
| MSFT   | Microsoft Corporation                  | 38,181        | 1,630            | 39,811  |
| NVDA   | NVIDIA Corporation                     | 35,282        | 1,502            | 36,784  |
| SPCE   | Virgin Galactic Holdings, Inc.         | 28,061        | 1,509            | 29,570  |
| FB     | Facebook, Inc.                         | 24,265        | 1,360            | 25,625  |
| DIS    | The Walt Disney Company                | 22,442        | 991              | 23,433  |
| BYND   | Beyond Meat, Inc.                      | 22,232        | 860              | 23,092  |
| NFLX   | Netflix, Inc.                          | 20,427        | 921              | 21,348  |
| JNUG   | Direxion Daily Jr Gld Mnrs Bull 3X ETF | 15,323        | 1,072            | 16,395  |
| GE     | General Electric Company               | 15,097        | 890              | 15,987  |
| RAD    | Rite Aid Corporation                   | 14,623        | 828              | 15,451  |
| SQ     | Square, Inc.                           | 12,866        | 788              | 13,654  |
| JD     | JD.com, Inc.                           | 11,529        | 585              | 12,114  |
| USO    | United States Oil                      | 11,263        | 581              | 11,844  |

Conventionally, submissions or comments that mention a ticker will spell it using uppercase letters, or following a dollar sign. However, a challenge is that not all uppercase words are valid tickers. A first match is made by identifying any succession of two to five capital letters. Subsequently, we used a pre-determined list, scraped from Yahoo Finance and Compustat, to check whether a match is indeed present in the available financial data. Some abbreviations or capitalised words which are not valid tickers might still show up, such as ‘USD’ (*ProShares Ultra Semiconductors*), ‘CEO’ (*CNOOC Limited*), and ‘ALL’ (*The Allstate Corporation*). Single characters also appear, such as ‘A’ (*Agilent Technologies, Inc.*). We manually created list of such tickers, and ignored featured matches, to build a preliminary list of candidates. We refined a second list of candidates by checking whether a collection of one to five letters, lower or uppercase, is preceded by a dollar sign. Any mention of ‘\$CEO’ or ‘\$a’ counts as ‘CEO’ and ‘A’, respectively. We checked these extracts again against the scraped list of available tickers.

### A.2 BERT Performance

We trained BERT on 90% of the labelled data, and used the remaining 10% for validation. Figure 8 plots two confusion matrices, one to test the performance of BERT out-of-sample, then in-sample. For the out-of-sample test, in Figure 8a, we only train BERT on 80% of the available data, then plot what the algorithm predicts for the remaining 20% of data. While BERT is good at differentiating



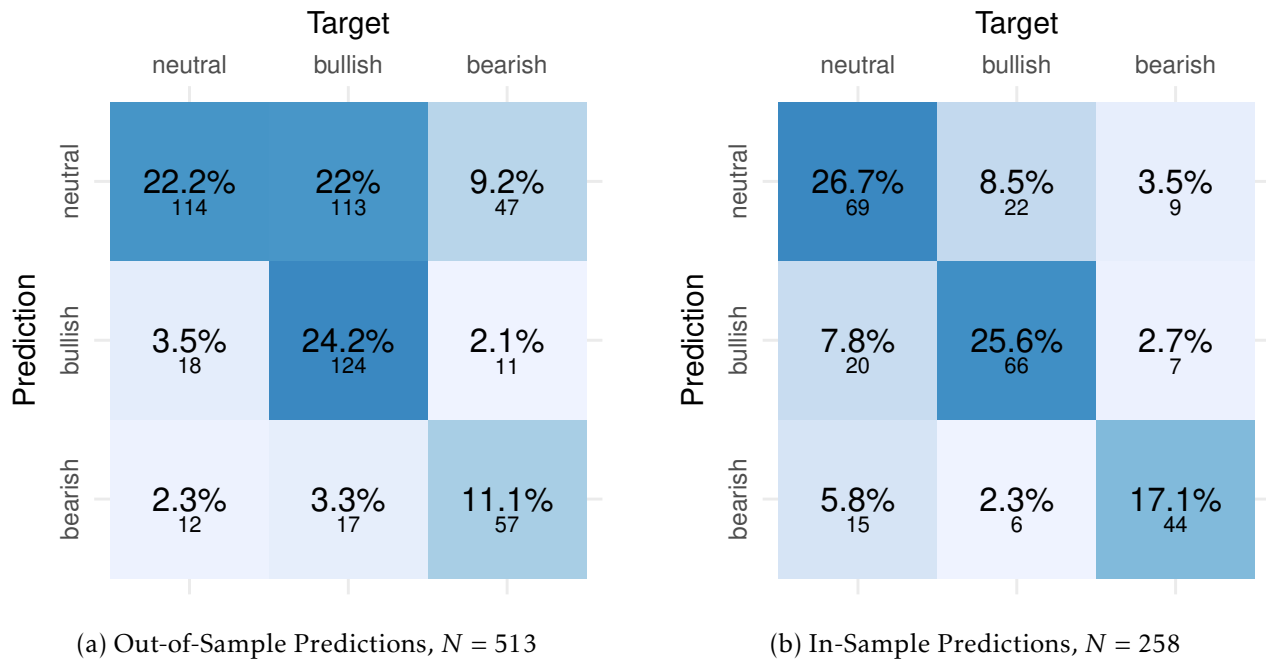


Figure 8: **Confusion Matrices for WSB Sentiments using BERT:** On the right, BERT is trained 80% of training data, and the performance is tested on the remaining 20% of training data; the algorithm performs well in differentiating between neutral and bearish submissions, but not so well in identifying bullish submissions. On the left, BERT is trained on all available data, and the performance is checked on the 10% of observations used for validation. Here, the algorithm performs better at identifying bullish submissions.

between neutral and bearish submissions, the classification of bullish sentiments suffers in this instance. However, when all the data is made available, the in-sample predictions, plotted in Figure 8b, are much better. This is likely a combination of a larger training set, as well as the data being used for validation against alternative models. The results are satisfactory; BERT correctly classifies 69% of ticker submissions in the validation set as neutral, bullish or bearish. This is better than a LASSO regression's accuracy, which we do not cover here.

### A.3 Data Appendix

Besides the scraped text data from Reddit, we downloaded historical price series on extracted tickers from The Center for Research on Security Prices (CRSP). The time period, denoted by  $t$ , is defined as the 24 hours between closing times on the New York Stock Exchange, namely 4pm EST. For example, a submission made on Tuesday at 5pm EST was categorised in the same time frame as a submission made on the following Wednesday at 10am. This is to mitigate any influence news occurring outside market hours might have on conversations in WSB. The choice of timing may not be consistent with non-US stock markets, but, given that almost all discussed assets are US stocks and indices, this does not affect the results. In all, the financial data used are  $P_{i,t}$ , the dollar price of security  $i$  at closing time of day  $t$ ,  $R_t^f$ , the risk-free rate of return obtained from Kenneth R. French's data library,  $(P_{i,t}^H, P_{i,t}^L)$ , the intra-day high and low prices of security  $i$  at closing time of day  $t$ , and  $V_{i,t}$ , the trading volume of security  $i$  at closing time of day  $t$ .

We computed the variables for Table 1 as follows:

- $R_{i,t}^e = \mathbb{E} \left[ \frac{P_{i,d \in t}}{P_{i,d-1}} - (1 + R_{d \in t}^f) \right]$ , where  $d$  is a day in calendar week  $t$ ,  $P_{i,d}$  is the closing price of asset  $i$  on day  $d$ , and  $R_d^f$  is the daily risk free rate on day  $d$ , and  $\mathbb{E}$  computes the arithmetic mean of price returns on each day  $d$  in week  $t$ ,
- $\sigma_{i,t}^e = \sqrt{\text{VAR} \left[ \frac{P_{i,d \in t}}{P_{i,d-1}} - (1 + R_{d \in t}^f) \right]}$ , where  $d$  is a day in calendar week  $t$ , and  $\text{VAR}$  computes the variance of price returns on each day  $d$  in week  $t$ ,
- $Q_{i,t} = \mathbb{E} [P_{i,d \in t} V_{i,d \in t}]$ , where  $V_{i,d}$  is the trading volume, in billions, of asset  $i$  on day  $d$ .

We calculated the variables used in Table 2 as follows:

- $R_{i,t}^e = \frac{P_{i,t}}{P_{i,t-1}} - (1 + R_t^f)$ , where  $P_{i,t}$  is the closing price of asset  $i$  on day  $t$ , and  $R_t^f$  is the daily risk free rate on day  $t$ ,
- $\overline{R}_{i,t}^e = \frac{P_{i,t-1}}{P_{i,t-6}} - (1 + R_{t-6}^f)^5$ , where the 5-period lag is based on a 5-day trading week,
- $\sigma_{i,t} = \sqrt{\frac{1}{10} \sum_{k=1}^{11} (R_{i,t-k}^e - \overline{R}_{i,t}^e)^2}$ , where  $\overline{R}_{i,t}^e$  denotes the average excess return observed between days  $t$  and  $t-10$ ,
- $W$  is the normalised submission adjacency matrix, where  $w_{i,j}$  is equal to the number of comments made by  $i$ 's author on an older submission  $j$  with the same extracted ticker, divided by the total number of comments made by  $i$ 's author on older submissions with the same extracted ticker,
- $a^s$  denotes a dummy variable for a submission with sentiment  $s$ , where  $s$  can be bullish ( $s = +1$ ), bearish ( $s = -1$ ) or neutral ( $s = 0$ ),
- $a_{-1}^s$  denotes a dummy, as above, for the sentiment of the last post made by the author on the same extracted ticker, where  $a_{-1}^{NA}$  represents the lack of such a submission.

We took two extra cleaning steps. First, to account for stock splits and mergers, we use CRSP's cumulative factors to adjust prices and volumes for each security  $i$ . Second, we remove any instances where the price is negative, or the excess return is greater than 10.

## A.4 Matching on Observables

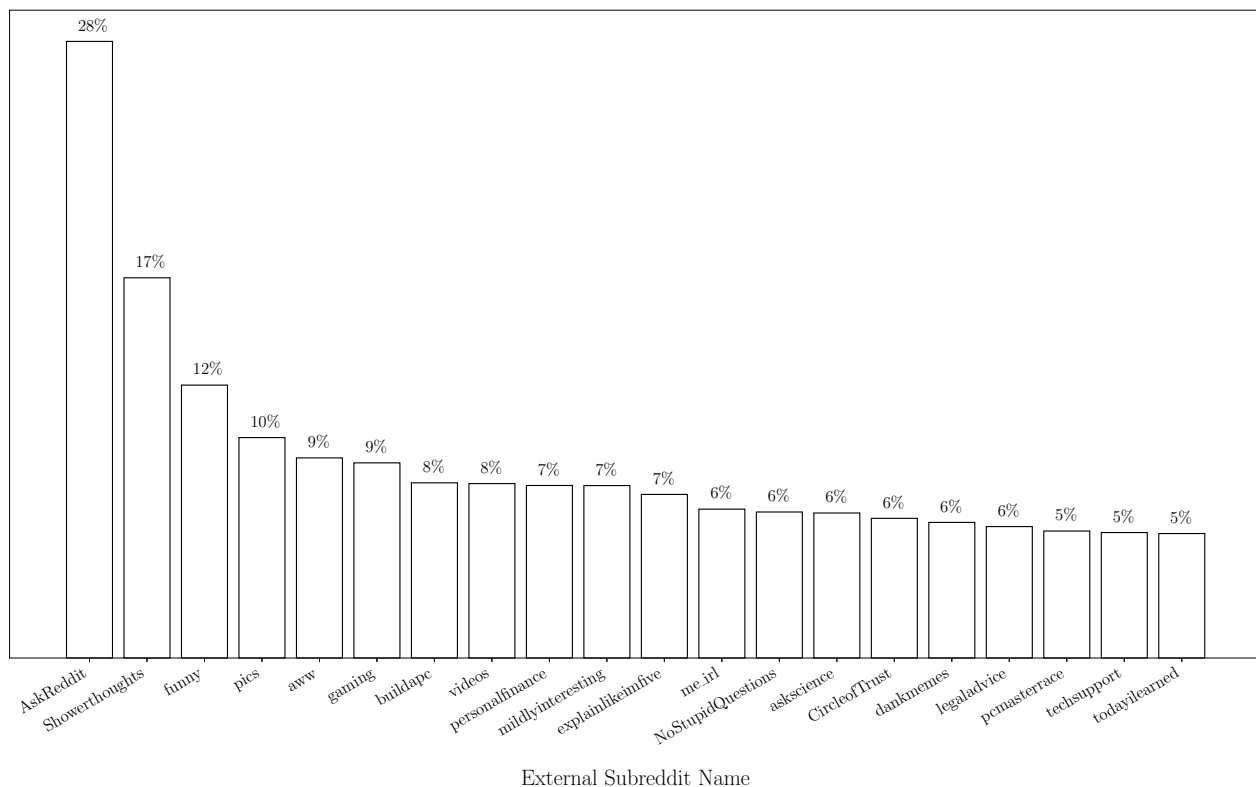


Figure 9: **Top 20 External Subreddits Frequented by WSB Users**; the plot shows the percent of WSB users who are externally active (posted on an external subreddit) who have posted on the subreddit labeled at the bottom of the chart.

**Distance Calculation and Matching** The matching on observables exercise attempts to isolate the impact that discussions on the WSB forum have on generating new discussions about the same asset. It attempts to isolate the "endogenous" component for asset interest. In order to accomplish this, it is important to control for several factors:

- i Whether users have the same activity profile on the forum: a user who is extremely active (routinely writes many comments and posts) may be more likely write a post about an asset than one who is less engaged
- ii Whether users are exposed to the same market environment: a user may post about an asset due to its market activity rather than because of discussions on the forum
- iii A user's "choice" to comment on one post versus another driven by previous asset interests

In order to control for these, we develop distance metrics to match users who comment on a post about a certain stock with those who do not, but exhibit similar characteristics otherwise - the more similar two users are in a specific fact, the lower their distance metric would be. Distance metrics are developed based on the following criteria to tackle the issues presented above:

- i Whether two users are active on the WSB forum for approximately the same time period and whether they have similar commenting or posting characteristics in the forum

- ii How shortly after a post about is made are users active on the forum: if two users are active on the forum at approximately the same point in time, they would be exposed to a similar market environment at the time of their activity
- iii Similar interests external to the WSB subreddit: certain user adopt more targeted interests and are active on forums such as “weedstocks” or “TeslaMotors”. These may correlate to which assets they are prone to be interested in. A plot of the top 20 external subreddits ranked by the number of unique WallStreetBets users who have posted on them, is presented for the interested user in Figure 9. It is important to note that this analysis uses external posts to compute similarity metrics. A more comprehensive analysis of external interest can be done by combining external comment and post data, however, commenting data across Reddit has a prohibitively large size for analysis at this stage. In order to compare user’s external activity, we match users based on the 20 subreddits they are most active in during the month of the submission they comment on. Several studies show that users change subreddit interests over time and, for this reason, we choose to match on external interests at a specific point in time.

In practice, this translates to the following matching exercise:

1. We matched individuals from the treatment group to all Redditors active (who have either posted or commented) on WSB (excluding those in the treatment group) who became active on the forum during the same month
2. We filtered out individuals in the potential control group whose last activity (post or comment) on the forum occurred before the submission that was commented on by their match in the treatment group (filtering out individuals who have become inactive before our period of interest finishes)
3. We filtered out individuals who create over 500 posts in a month as we discovered most such users to be bots; this only eliminate 0.015% of WSB users
4. Lastly, we filtered out individuals in the potential control group who have posted about a ticker before the time of the relevant submission. This indicates that the individual is *already* interested in the ticker and the events surrounding our submission of interest would have no effect
5. We then calculated the behavioural distance between a ticker commenter and the remaining potential matched control group using several distance metrics:

$D_1$ : the time between the submission the commenter interacted with and the next activity of the member of the control group. One of the main sources of homophily in this study is the appearance of news or market moves by a particular ticker. Section 2.2 shows a strong correlation between asset returns sentiment and posting volumes. We controlled for this by making sure that a member of the control group was active shortly after the time of the relevant submission in order to have been exposed to the same market and news environment at the time of their activity

$$D_1 = \min_{t_{i,j} > s_k} [t_{i,j} - s_k, 30 \text{ days}] \quad (19)$$

where  $t_{i,k} \in$  times of activity of  $j$  (a member of control group),  $s_k$  is the time of the submission a member of the treatment group,  $k$ , commented on. We add 30 days to the minimization because if an

individual is active after over 30 days from a given post, the market will have changed substantially and they are likely not exposed to the same market events

$D_2$ : the average comment and submission length on WSB

$D_3$ : the average comment and submission amount on WSB

$D_4$ : the difference in the number of external subreddits that both individuals posts on (a measure of how many external interests each individual has)

$D_5$ : the overlap, in percentage terms, in the number of external subreddits the two individuals are both interested

$D_6$ : the difference in the average number of posts each individual makes on external subreddits (a measure of how active, on average, each individual is within external communities)

6. We normalized all distance variables  $D_i$  between  $[0,1]$  using min-max normalization, summed them ( $D = D_1 + D_2 + D_3 + D_4 + D_5 + D_6$ ) and computed their inverse ( $1/D$ ) to use this as our final metric  $M_{i,j}$  between a member of the control,  $i$ , and member of the treatment group,  $j$ . A small  $M$  would imply that two individuals are a poor match, a large  $M$  implies that they are a good match.

If the individual comments on multiple submissions, the  $D_1$  metric was computed separately for each submission the individual comments on in order to control for potential exposure to each market event. For individuals who commented on over five different submissions, we considered only the most recent five submissions for the purposes of the distance calculation, in order to proxy recent activity and market events which may have prompted asset interest. For someone who comments on four submissions, the distance calculation between them and a member of the control group would be computed as:  $D = D_{1_1} + D_{1_2} + D_{1_3} + D_{1_4} + D_2 + D_3 + D_4 + D_5$ .

7. Lastly, we solved the following maximal-matching optimization problem in order to match members of the treatment and control groups:

$$\begin{aligned}
& \max \sum_{i,j \in N \times T} M_{i,j} x_{i,j}, \\
& s.t. \sum_j x_{i,j} \leq 1 \quad \forall i \in N, \\
& \sum_i x_{i,j} \leq 1 \quad \forall j \in T, \\
& 0 \leq x_{i,j} \leq 1 \quad \forall N, T, \\
& x_{i,j} \in \mathbb{Z} \quad \forall N, T,
\end{aligned} \tag{20}$$

where  $N$  stands for our control group,  $T$  stands for our treatment group.

Practically, even though this is an integer program, we can drop the last integrality constraint.

A question that might arise is what the distribution in specific distance values ( $D_1, \dots, D_6$ ) is, and whether, even though over distance  $D$  is minimized, the individual distances can take on large values. We examine the max-min normalized distributions and observe that, even though it is possible for individual distances to take on large values, empirically this occurs rarely as there are a sufficient number of potential users to match with to ensure all criteria are well-matched against. Empirical details are available upon request.

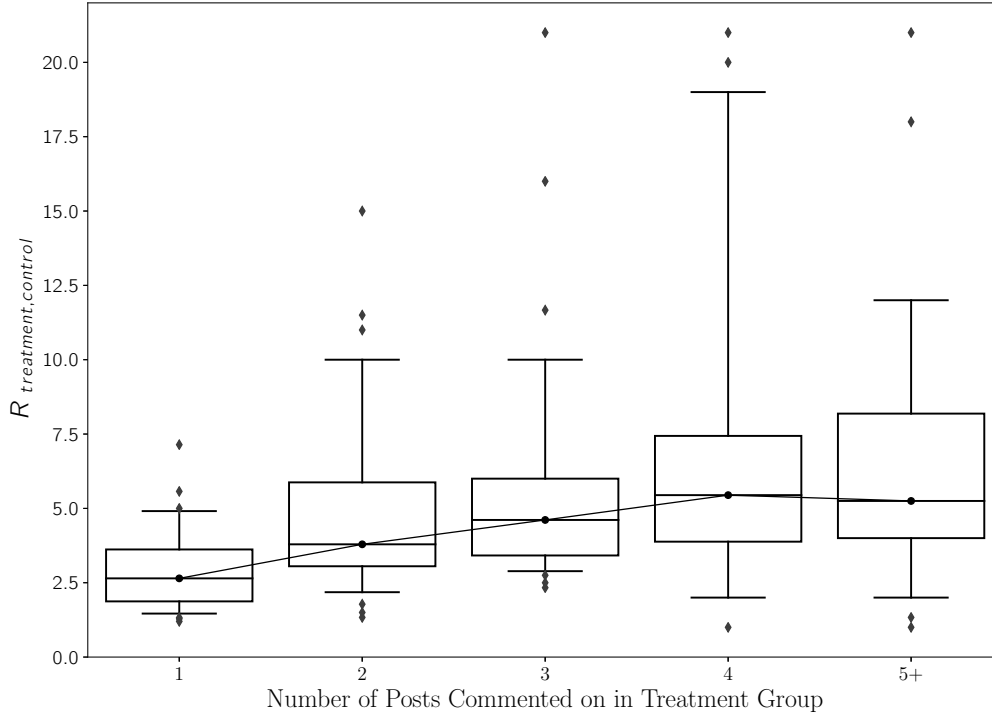


Figure 10: **Estimated Effect of Commenting on Posting a New Submission**; the distribution of the estimated treatment effect,  $R_{treatment,control} = \hat{\pi}_j^T / \hat{\pi}_j^C$ , across the top 1% of tickers is graphed as a series of box plots. The solid line connects the median estimate, and the protruding dashes represent the 5<sup>th</sup> and 95<sup>th</sup> percentiles.

**Evaluating Impact of Commenting** The effect of commenting on a post on future probability to create a post, displayed in Figure 6, appears small in absolute terms because there is one order of magnitude more comments than submissions. Figure 10 presents the ratio of proportion,  $\hat{\pi}_j^T / \hat{\pi}_j^C$ , estimates as a function of the frequency with which a user has commented. The first important result highlighted is that all differences in proportion are positive, implying that engaging in discussions on WSB about an asset strictly increases future interest in the asset. This effect becomes more prominent the more a user comments on ticker-related posts, evidencing some form of threshold contagion in investor interest spread. For this portion of the analysis, we filter out instances when no members of the control group create a post in a given treatment category as this renders the ratio of proportions equal to infinity.

## A.5 Model

**Longer Time Scales and Elasticities of Demand** On longer scales, larger bull runs form, depending on the responsiveness of smart and ordinary investors to price changes. To see this, write  $Y(p)$ , the demand function for ordinary investors, and  $S(p)$  for ordinary investors. This time, we assume

$$\frac{\partial S(p)}{\partial t} = 0,$$

so that the adjustment in shares held by smart investors is purely endogenous with respect to price. The total derivative of Eq. 4 thus yields

$$\frac{\dot{p}}{p} = \frac{1-s}{(\varepsilon_Y + s[\varepsilon_S - \varepsilon_Y])} \frac{\dot{Y}}{Y}, \quad (21)$$

where

$$\varepsilon_Y = -\frac{p}{Y} \frac{\partial Y(p)}{\partial p}$$

is the price elasticity of demand for ordinary investors, and, likewise,

$$\varepsilon_S = -\frac{p}{S} \frac{\partial S(p)}{\partial p}$$

is the price elasticity of demand for smart investors. We denote  $s = S/Q$ , the ratio of shares owned by smart investors. Eq. 21 demonstrates that prices may shoot off indefinitely if, between all investors, increasing prices fail to trigger enough shareholders to sell. A crash would only take place once enough investors decide to increase their response to higher prices, causing a sign reversal in Eq. 21, and a subsequent rout. This type of model, where the responsiveness of agents' demand to price varies, is studied extensively by Hommes (2013), and so we do not dwell on it here.

**Herding and Investor Clusters** This paper offers a deeper exposition of asset demand by ordinary investors by introducing the effect of consensus among ordinary investors, and its consequences on price stability. To that end, the particular appeal of social dynamics is in producing movements large enough to overcome the market depth bottleneck, detailed in Eq. 4 in its simplest form. To see this, consider the breakdown of small investors  $i$ , which we replicate from Bouchaud & Potters (2003):

$$\Phi = \sum_i \phi_i,$$

where  $\Phi$  replaces  $Y$  as total demand, to reflect individual demand  $\phi_i$  taking values  $-1$  should  $i$  sell,  $0$  if they are neutral, and  $+1$  if they buy. There is an implicit assumption that all investors  $i$  have the same purchasing power, which we will denote  $M$ ; while unrealistic, we make this simplifying assumption to highlight the impact of consensus. From Eq 5, it is obvious that individual ordinary investors, controlling an infinitesimally small fraction of the supply  $Y$ , are unable to translate their trades into large price movements. However, should investors be linked, such that they coordinate their demand simultaneously, then

$$\Phi = \sum_C L(C) \phi_C,$$

whereby ordinary investors are organised into  $C$  clusters of size  $L(C)$ , trading in direction  $\phi_C$ . This is where our simplifying assumption is useful: heterogeneity in investor size implies that certain clusters emerge from large investors, confusing the effect of consensus by which many small investors  $i$  can form an equally-sized clusters. Therefore, this formulation highlights social dynamics, i.e. interactions *between* individuals, as the mechanism for ordinary investors to drive market fluctuations.

**Derivation of Eq. 13** In this section, we show the steps between Eq. 12 and Eq. 13, which is

$$\frac{\dot{p}}{p} = \frac{M}{pQ} \times \frac{d}{dt}(A\phi).$$

The derivative with respect to time is

$$\frac{\dot{p}}{p} = \frac{M}{pQ} \left( \frac{dA}{dt} \phi + \left[ \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial p} \frac{dp}{dt} \right] A \right),$$

where we take the partial derivative of  $A$ , and break down the derivative of  $\phi$  with respect to time and price, since both determine the overall change in  $\phi$ , as discussed in section 4.3). Rearranging yields

$$\frac{\dot{p}}{p} - \frac{MA}{pQ} \left( \frac{\partial \phi}{\partial p} \frac{dp}{dt} \right) = \frac{M}{pQ} \left( \phi \frac{dA}{dt} + A \frac{\partial \phi}{\partial t} \right),$$

and the left-hand side is simplified as

$$\frac{\dot{p}}{p} \left( 1 - \frac{MA}{Q} \left( \frac{\partial \phi}{\partial p} \right) \right) = \frac{M}{pQ} \left( \phi \frac{dA}{dt} + A \frac{\partial \phi}{\partial t} \right),$$

given  $\dot{p} = dp/dt$ . Substituting  $dA/dt = cAV - rA$  from Eq. 6 into the right hand side yields

$$\frac{\dot{p}}{p} \left( 1 - \frac{MA}{Q} \left( \frac{\partial \phi}{\partial p} \right) \right) = \frac{M}{pQ} \left( \phi [cAV - rA] + A \frac{\partial \phi}{\partial t} \right).$$

Factoring  $A$  on the right-hand side of the equation and noting  $\dot{\phi} = \partial \phi / \partial t$  yields

$$\frac{\dot{p}}{p} \left( 1 - \frac{MA}{Q} \left( \frac{\partial \phi}{\partial p} \right) \right) = \frac{MA}{pQ} (\phi [cV - r] + \dot{\phi}). \quad (22)$$

which gives us the components (1), (3), (4) on the right hand side of our equation as outlined in section 4.3. Recall that

$$pQ = pY + pS,$$

and denote  $s = S/Q$ , the proportion of shares owned by smart-money investors. Finally, the ordinary investor asset demand identity holds,  $pY = \phi MA$ , such that the total value of shares owned by ordinary investors ( $pY$ ) equals the number of active ordinary investors in the market,  $A$ , the average amount of money they hold,  $M$ , and their buying intensity  $\phi$ . Thus, we rearrange the the identity

$$\begin{aligned} pQ &= pY + pS, \\ 1 &= \frac{\phi MA}{pQ} + s, \\ (1 - s) &= \frac{\phi MA}{pQ}. \end{aligned}$$

We define the price elasticity of buying intensity as

$$\varepsilon_{\phi} = -\frac{p}{\phi} \frac{\partial \phi}{\partial p},$$

such that

$$\begin{aligned} (1 - s)\varepsilon_{\phi} &= -\frac{\phi MA}{pQ} \times \frac{p}{\phi} \frac{\partial \phi}{\partial p} \\ &= -\frac{MA}{Q} \frac{\partial \phi}{\partial p}. \end{aligned}$$



Substituting this results into Eq. 22 yields

$$\frac{\dot{p}}{p} \left( 1 + (1-s)\varepsilon_\phi \right) = \frac{MA}{pQ} \left( \phi[cV - r] + \dot{\phi} \right).$$

Eq. 12 is obtained by dividing by the elasticity term:

$$\frac{\dot{p}}{p} = \frac{MA}{pQ} \times \frac{\phi(cV - r) + \dot{\phi}}{1 + (1-s)\varepsilon_\phi}.$$

## A.6 Asset Price Simulation

As we discussion in Section 4.3, our model can reproduce a diverse set of patterns. In this section, we consider the impact of perturbing our initial parameters:  $\alpha$ ,  $\lambda$ ,  $a$  and  $b$ .

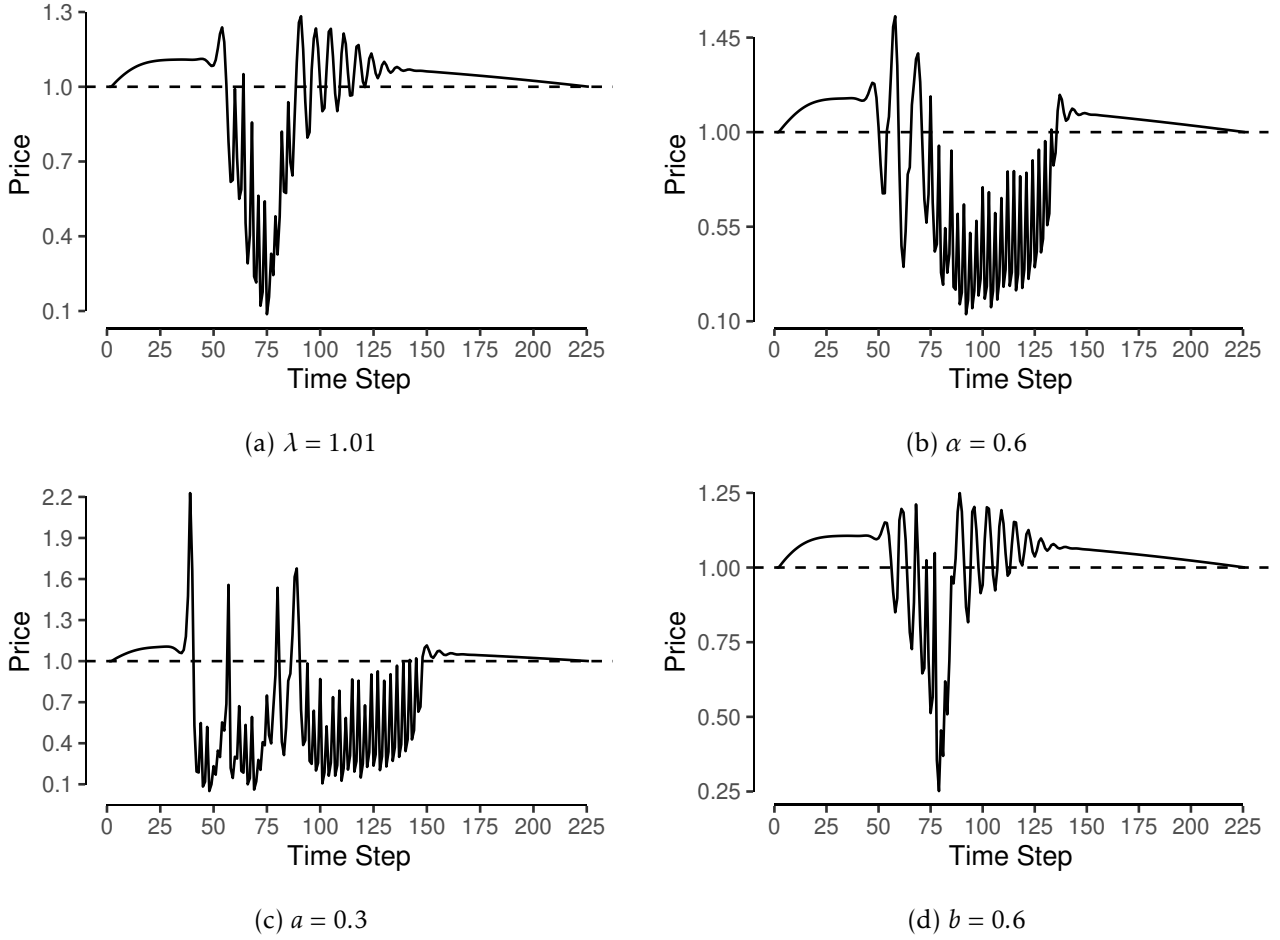


Figure 11: **Price Dynamics**; these four plots demonstrate the qualitative changes in the price dynamics after slight alternations of each parameter. The baseline parameters are  $\lambda = 0.99$ ,  $\alpha = 0.45$ ,  $a = 0.2$ ,  $b = 0.5$ . The alternations in each figure is noted in the caption.

Figures in 11 reproduce Figure 7 with slight changes to demand parameters. In all cases, the initial price increase remains. The qualitative results are broadly summarised as follows: i) an increase in coordination through  $\lambda$  produces a sharper crash, and faster recovery, ii) higher demand persistence  $\alpha$  produces larger oscillations during the crash, iii) more trend following  $a$  produces bigger, sporadic price spikes, which are more sensitive to price *acceleration*, and iv) higher aversion to volatility  $b$  produces a sharper crash, but faster and stable recovery.

## A.7 Contagion Model

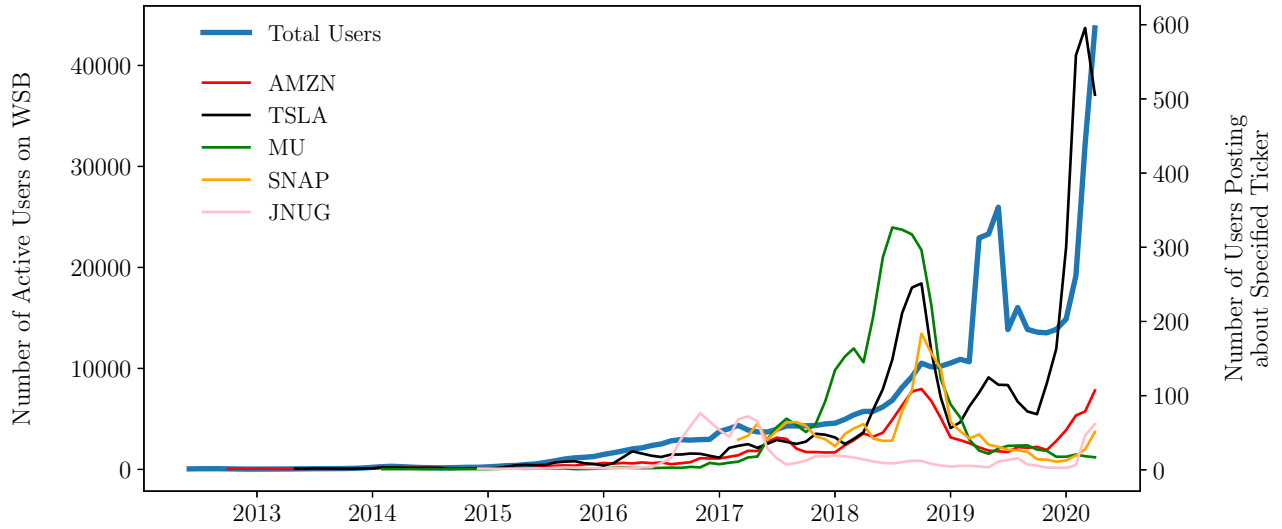


Figure 12: **Ticker Activity versus Overall WSB Activity**; the number of unique monthly active users (users who comment or post) on the left y-axis plotted against the unique number of posters about a specific ticker (on the right y-axis). Both are plotted as 3-month rolling means.

**Empirical evidence for the contagion model** We provide evidence that discussions about assets are self-perpetuating on WSB: social contagion appears to play a role in the assets that investors discuss and consider for their portfolios. Here, we attempt to reconcile the size of discussions on WSB over time with our contagion model of choice.

We observe that some of the most commonly mentioned tickers appear to exhibit a hump-shaped pattern of number of interested users over time. Figure 12 shows the number of authors interested in a given ticker, aggregated on a monthly basis, versus the overall number of unique users active on WSB. Unique ticker interest shows a more ‘bursty’ and hump-shaped pattern which, at times, is distinct from the overall trend of number of users active on the forum. We use data from WSB to estimate our contagion parameters,  $c$  and  $r$ , at different points in time and show that our simple contagion model captures the hump-shaped ticker interest. One challenge is that our epidemic model is well-suited to model a single “outbreak” in ticker interest as it assumes a static population. We, therefore, look at how well our model fits data with non-overlapping periods of discussion intensity. Appendix A.7 offers insight into how parameters are initialized and into model sensitivity.

Figure 13 shows the *predicted* discussion sizes of specific tickers from the contagion model (as a percent of active users on the forum at the time) versus the *actual* sizes of ticker discussions on the WSB forum. Despite its clear shortcomings, this simple model that is tuned on noisy global parameters still captures the overall shape of social contagion in ticker interest. It exhibits some desirable properties, such as the fact that contagion spreads more slowly in communities that are sparse and faster as the community grows: consider Figure 13a where JNUG interest gradually permeated the forum over the course of a year, whereas TSLA interest in 2020, displayed in Figure 13d, spread to a much greater population over the course of three months.

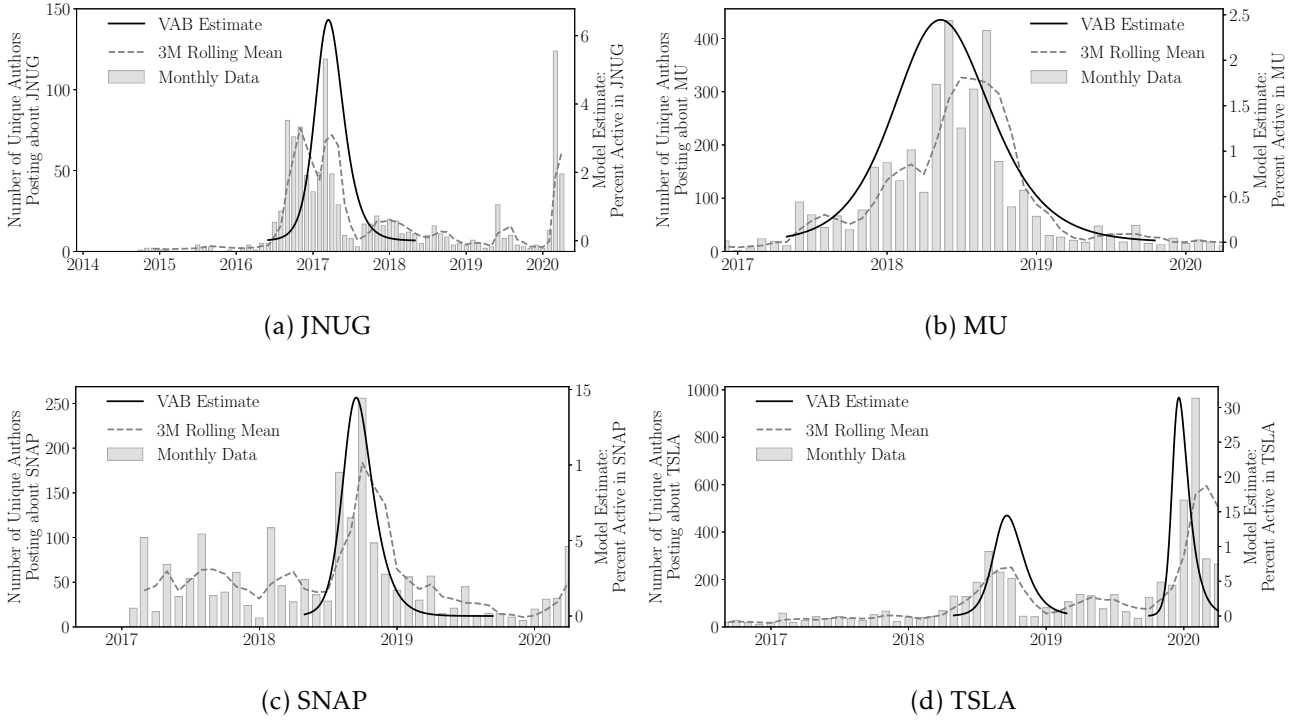


Figure 13: **Contagion Model Forecasts for Active Investors in Assets on the WSB Forum**; the contagion model is initialized with the parameters at the start of an “outbreak” of activity on WSB in a specific ticker and used to approximate the growing interest and, subsequent, loss of interest.

**Estimating  $c$  and  $r$  Parameters Globally** Certain contagion variables are more straightforward to estimate from WSB data than others. For the moment, we assume the total population at a given time,  $N_t$ , to be the number of active users on the forum: those who comment or post at time  $t$ . The number of users initially interested in asset  $i$ ,  $A_{i,0}$ , is estimated as the number of people who post about  $i$  in some starting period 0. The average dollar amount each user has at their disposal,  $M$ , can be tweaked from available surveys<sup>15</sup>. All of these are adjusted to check the results’ sensitivity.

Contagion parameters,  $c$  and  $r$ , require more care in isolating the effect of social interaction from other variables that may impact asset interest.

The contagion rate is comprised of two parameters:

$$c = \tau\gamma, \quad (23)$$

where  $\tau$  is transmissibility, the probability of transmission of asset interest given contact between a passive and active investor), and  $\gamma$  the average rate of contact between passive and active investors on the forum. The first,  $\tau$  is easily estimated from the results in Figure 6, showing a spread of ticker-specific interest at a rate of 1.1% in the single treatment group category. In order to estimate  $\gamma$ , we counted the total number of unique users who commented on a submission, on average (this is an approximation for the number of “susceptible” individuals that an infected person comes in contact with), multiplied by the number of posts that a user makes, on average, about a given ticker, and divided by the number of users active on the forum, measured as unique users who post a submission or comment, while the submissions about the ticker are discussed on the forum. We consider interactions to be uniform over the duration of the “infectious period” so divide this number by the

duration. The average contact rate is estimated at  $\gamma = 1.2 \times 10^{-4}\%$ . This may seem particularly small, however, intuitively this is not surprising as *gamma* captures how many people will pay attention to a single submission versus all the other content in a thriving online community. The contact rate also differs greatly from one month to the next as people join and leave the forum, so this rate is more precisely tracked. This returns a combined contagion rate of  $c = 9.8 \times 10^{-7}$ .

Estimating  $r$  presents a challenge: users do not advertise a loss of interest in an asset on WSB. In order to estimate the recovery rate, we turn to its classic definition as ‘the inverse of the infectious period,’ as described in Ridenhour et al. (2018). We consider the ‘infectious period’ the time difference between when an author make his first post about a ticker and the time when his last post ceases to receive comments (ceases to be actively discussed on the forum). During this period, an individual spreads their perspective about an asset on the forum. We estimate the recovery rate at approximately  $r = 1/19.6 = 5.1 \times 10^{-2}$ . Further details on estimates for contagion parameters and the sensitivity of the model are presented in the following section.

**Initializing  $c$ ,  $r$ ,  $N_0$  and  $A_0$  for Social Contagion in a Specific Ticker** We initialize  $N_0$  as the total number of unique authors who are active on the forum from the start to 5 months after the beginning of the contagion: we select this period in order to estimate the size of the population who would be affected by the excitement about a ticker, but we make the period sufficiently short so as to avoid capturing inflated population numbers driven by user churn (when a group of users becomes uninterested in a given subreddit and is replaced by a new group of users). We initiate the  $A_0$ , the initial population that begins to discuss a ticker on the forum driving further interest, as the average number of of unique authors posting about the ticker in the 3 months preceding the uptick. We also adjust gamma to be calculated at the time the uptick begins.

**Units and Further Discussion on Initialization of Parameters** We check our contagion units with the standard epidemiology literature and intuition. The change equations in Eq. 6 should result in a unit of:  $\frac{\text{persons}}{\text{day}}$ . We consider our input into the  $c$  constant; we know that  $A, V$  have a units of  $\text{persons} \times \text{persons}$ , therefore our  $c$  constant must have units of  $\frac{1}{\text{person-day}}$ . Our  $c$  constant is composed of  $\tau$  and  $\gamma$ .  $\gamma$  helps us quantify how many contacts occur between *passive* and *active* individuals each day;  $\tau$  tells us what proportion of the *passive* investors become *active* as a result.  $\tau$  is a unitless constant as it is in persons who become *active* divided by persons contacted in total, or  $\frac{\text{persons}}{\text{persons}}$ . The  $\gamma \times AV$  should give us the total contacts between *passive* and *active* per day in units  $\frac{\text{person}}{\text{day}}$ .  $\gamma$  should give us, for every *active* individual, on a given day they are active, what is the probability that a *passive* person comes in contact with them (in units of 1 divided by person  $\times$  day)? Subsequently,  $\gamma \times A$  would give us the number of *active* investors that a single *passive* investor comes into contact with each day. Finally,  $\gamma \times AV$  would give us the total contacts between the two. We, therefore approximate  $\gamma$  with this in mind. The  $\gamma$  parameter is sometimes referred to as the fraction of *active* individuals that a passive individual comes in contact with. However, on the WSB forum, the fraction of *active* individuals is relatively small compared to the overall population. We therefore estimate the fraction of active individuals as the probability to come in contact with a single active individual (essentially discounting the probability that a *passive* individual will encounter multiple *active* individuals in a given day). This decision is justified both given how few *active* individuals there are as compared to the overall size of the forum and given that in Figure 7a we show that 71% of people interact with one post before posting themselves. (or becoming *active*). We consider, over the duration of an *active* individual posting on the forum, what fraction of the *passive* population come in contact with this individual. We divide this by the total number of days that the activity lasts (to add a time

dimension).

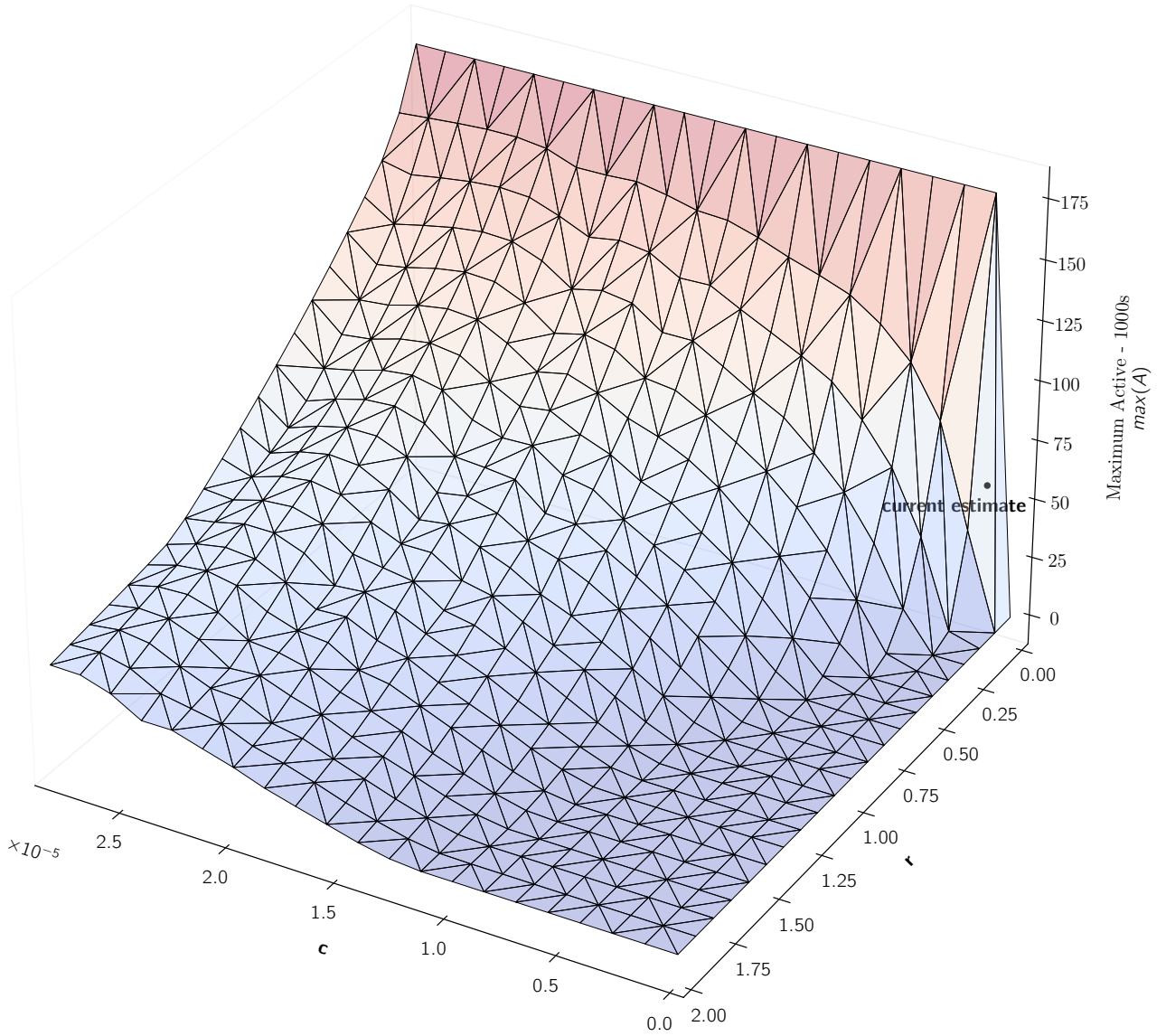


Figure 14: **Sensitivity of  $\max(A)$  to  $c$  and  $r$  Parameters**; this plot shows the dependence of our  $\max(A)$  on various initialization parameters. This plot highlights that the current estimate lies at a point with a large gradient where  $\max(A)$  is not 0 but also not the entire population. This is consistent with empirical observations where not all investors become interested in a given asset - even for assets which are most popular on the forum.

**Sensitivity of the Maximum of Active Investors to  $c, r$**  In our model, we are most interest in the number of *active* investors,  $A$ , at a given point in time. This quantity approximates how many retail investors that are susceptible to social contagion are actively interested and investing in a given asset at a point in time. Specifically, we are interested in the  $\max(A)$  over the course of an entire “social epidemic” as this would be tied to the maximum market impact that the social contagion would have. We, therefore consider how  $\max(A)$  varies with our initial  $c$  and  $r$  parameters. We initialize our model with the parameters in our second Tesla discussion, plotted in Figure 13d and observe how  $\max(A)$  changes in the surrounding regions for  $c$  and  $r$ , plotted in Figure 14. We observe that as  $r$  increases (which is the inverse of the duration for which an investor remains interested in a given asset) the  $\max(A)$  decreases: as people loose interest quickly, they have less time to spread

their interest to others and, therefore, the excitement in the given asset dies out with very few active investors at a given point in time. We also observe that as  $c$  increases, which is our contagion rate, the number of active investors at a point in time increases. A final important point is that a fairly rough approximation for the  $c$  and  $r$  parameters from our data place the estimate in a “critical region” of the graph: a region with a high gradient - a region where not all investors on the forum become active in the specified asset. This is intuitive as we do not observe that all investors discuss a single asset; rather a variety of assets are always discussed with different investors becoming part of different ticker discussions.

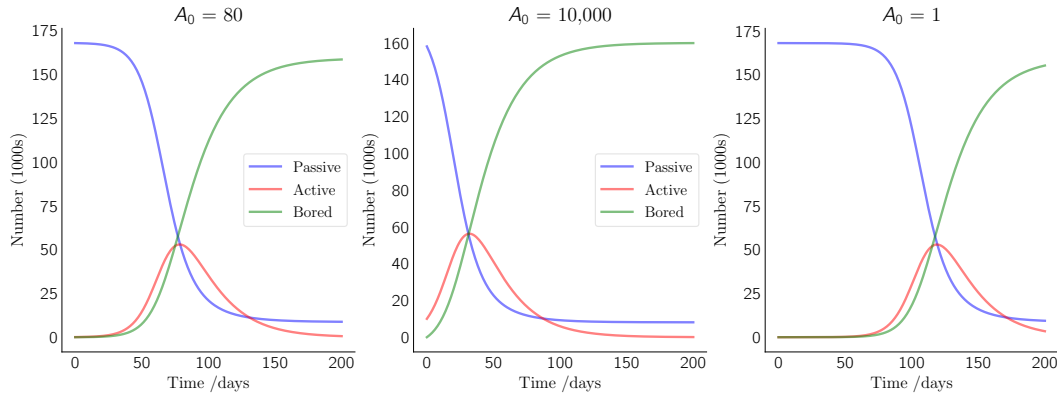


Figure 15: **Sensitivity of Investor Dynamics to  $A_0$** ; this plot shows that the dynamics are relatively insensitive to how we initialize  $A_0$ .

**Sensitivity of Contagion Dynamics to  $A_0$**  We also consider how dynamics develop for different choices of  $A_0$ , the number of initial investors who become interested and spread interest in an asset, and plot the evolution for two different values of  $A_0$  in Figure 15 while keeping all other parameters constant. The model appears relatively insensitive to this parameter. A change of  $A_0$  from 80 to 10,000 drives only a modest change in the evolution of aggregate investor interest. Interest spreads faster in the latter, however,  $\max(A)$  appears relatively consistent demonstrating that other parameters contribute more significantly to  $\max(A)$  and the overall dynamics. The graph with  $A_0$  highlights a potential flaw with the model; when the right epidemic conditions are met, even a small number of *active* individuals can trigger a large surge of interest. However, our parameters are extracted from a ticker and time when social contagion in *TSLA* does occur. The spread of interest is not a given for all initialization parameters and, in fact, is relatively rare.

**Sensitivity of Contagion Dynamics to  $N_0$**  In WSB, we have a changing population (which we alter in our initialization for the model presented in Figure 13). We observe how dynamics develop for a different  $N_0$ , while keeping all other parameters constants, in Figure 16. Our model appears highly sensitive to this parameter. Even though in the end, the interest spreads to a total number of investors that is proportional to the increase in  $N_0$  (this is observed in the final number of *bored* individuals), the interest spreads much quickly for a large  $N_0$ . This is particularly important as this is tied to a larger  $\max(A)$  and a larger potential market impact generated by the discussion.

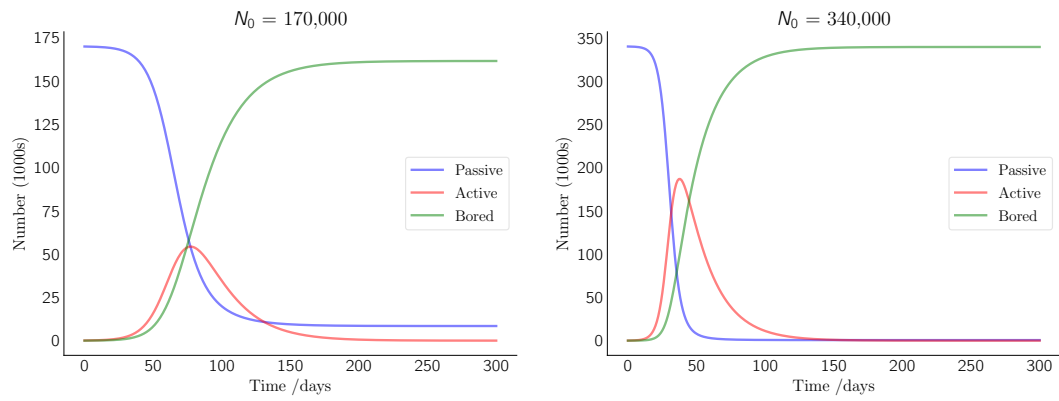


Figure 16: **Sensitivity of Investor Dynamics to  $N_0$** ; this plot shows that the dynamics are highly sensitive to how we initialize  $N_0$ . A larger  $N_0$  results in a disproportionate increase in  $\max(A_0)$  and a fast development of the dynamics.



## A.8 Topic Model

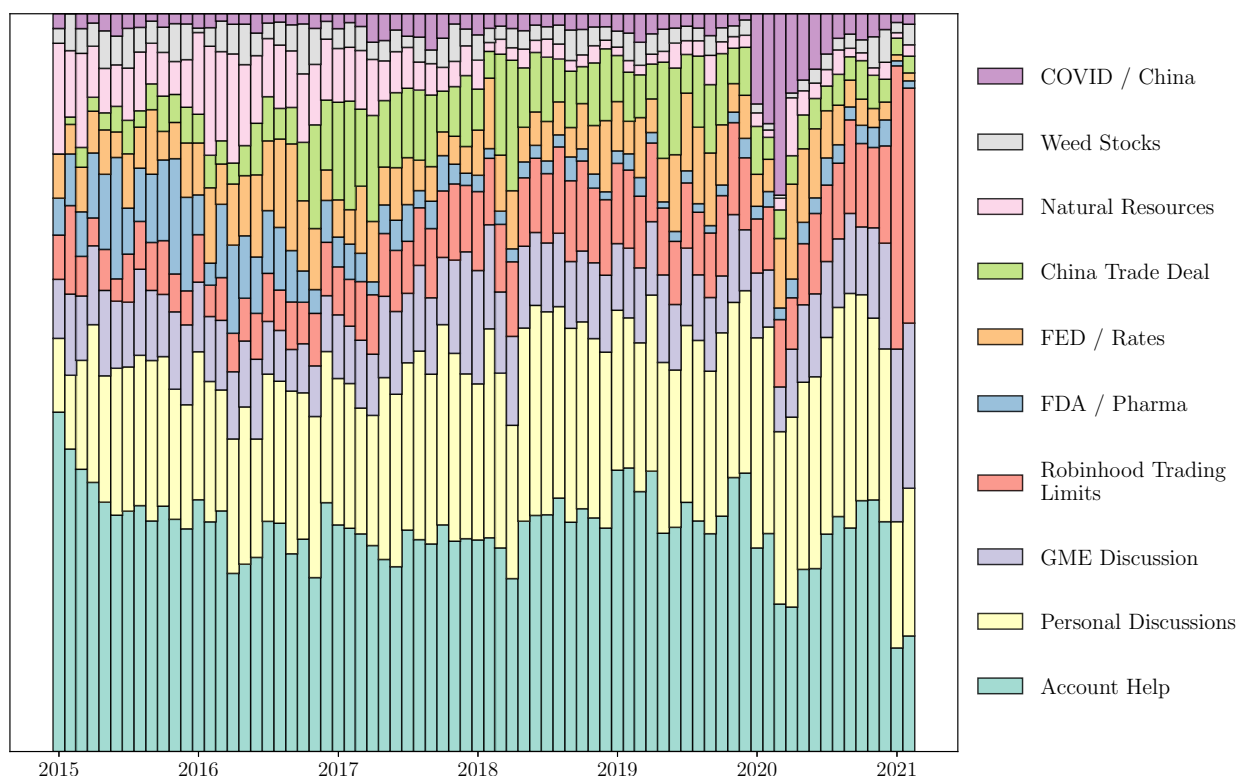


Figure 17: **Temporal Trends in Topics**; the stacked count, normalized to 100% at each time period, showing the prevalence of a select subset of topics discussed on WallStreetBets.

Does WSB reflect new information for the larger market to trade on, or social activity that drives perceived changes in value, regardless of fundamentals? A topic model offers a simple method to evaluate the content of WSB discussions. Figure 17 presents our preferred topic model, namely the Biterm Topic Model (BTM), which is optimal for smaller bodies of text (Yan et al. 2013). Submissions from April 2012 to February 2021 give a time series of almost 100 months. A random sample of submissions is drawn for the months of January and February 2021 in order to prevent these two months with a high number of submissions from skewing the topic model results.

Figure 17 presents a stacked plot of the monthly submission count of a selected subset of discussions, normalised by the total across the selected topics. It begins in 2015 when the forum gained a consistent user base. On one hand, some topics persist in the overall discussion: people consistently ask for advice about trading accounts and anonymously share details of how their trading is affecting their personal lives. On the other hand, topics concerned with larger economic trends wax and wane over the observation period. Two examples of this are the uptick in submissions discussing GME and Robinhood account trading limits, coinciding with the GME short squeeze, and the COVID-19 topic, which is negligible until January 2020, but gains prominence in the subsequent months. Pharmaceutical company and natural resource discussions, on the other hand, seem to lose popularity. A full list of topics with their respective keywords is presented below.

| Topic Title                 | Top Words  | Topic Prevalence (%) |
|-----------------------------|--|----------------------|
| Robinhood Trading Limits    | robinhood, gme, account, nkla, margin, order, app, limit, broker, orders           | 1.8                  |
| International Trade         | expected, yr, china, usd, europe, japan, pmi, manufacturing, korea, data           | 0.8                  |
| Retail Sales + Amazon       | sales, amazon, home, stores, business, online, companies, food, store, retail      | 3.0                  |
| Top Stock Picks / Positions | tsla, news, sold, aapl, weeks, holding, hold, amd, months, dip                     | 11.6                 |
| Other                       | comments, daily, best, moves, spy, weekend, fo, fn, fm, fp                         | 1.1                  |
| Other                       | mentions, vote, log, wsbvotebot, submission, posts, check, reverse, mention, great | 0.2                  |
| Electric Cars               | tsla, energy, car, ev, cars, nio, electric, battery, elon, space                   | 2.3                  |
| Revenues, Earnings, Ratings | revs, beats, tgt, line, eps, neutral, downgraded, initiated, fy, reports           | 0.9                  |
| FDA / Pharma                | drug, fda, phase, patients, vaccine, trial, treatment, results, clinical, covid    | 2.9                  |
| Revenues, Earnings, Ratings | revenue, million, growth, quarter, share, sales, billion, net, expected, eps       | 4.0                  |
| China Trade Deal            | trump, china, said, president, deal, bill, house, election, chinese, news          | 3.1                  |
| Social Media Stocks         | fb, game, snap, aapl, disney, games, video, google, users, netflix                 | 2.7                  |
| GME Discussion              | companies, gme, investors, world, years, believe, hedge, actually, value, funds    | 7.3                  |
| Financial News              | data, information, news, financial, report, based, find, sec, research, investors  | 4.3                  |
| Personal Discussions        | life, wife, ass, little, said, spy, getting, old, red, went                        | 7.6                  |
| Weed Stocks                 | million, share, capital, ceo, cannabis, ipo, public, management, merger, billion   | 2.1                  |

Table 8: Topics Extracted from BTM Model

#### Topic Titles and Top Words, by Mention, in each Topic

| Topic Title          | Top Words  | Topic Prevalence (%) |
|----------------------|--|----------------------|
| Software Tech Stocks | amd, aapl, intc, companies, data, cloud, software, technology, services, tech      | 3.4                  |
| Earnings Release     | release, estimates, consensus, share, revenue, move, open, beat, average, interest | 1.0                  |
| Natural Resources    | oil, gold, prices, silver, gas, futures, crude, production, demand, companies      | 1.8                  |
| FED / Rates          | fed, rates, rate, economy, markets, economic, said, interest, growth, inflation    | 6.0                  |
| Other Tech Stocks    | tsla, pltr, elon, musk, mods, ban, gme, retards, gains, autists                    | 2.0                  |
| COVID / China        | virus, covid, cases, coronavirus, china, world, weeks, corona, states, news        | 4.8                  |
| Other                | spy, bear, gay, text, bears, bull, gang, msft, words, stonks                       | 1.7                  |
| Debt / Loans         | debt, cash, pay, credit, loans, loan, million, interest, billion, bank             | 3.4                  |
| Other                | usd, bln, exp, revenue, newswires, eps, co, share, symbol, live                    | 1.4                  |
| Other                | spy, chart, close, index, month, performance, major, past, futures, sectors        | 1.3                  |
| Account Help         | account, help, investing, best, start, robinhood, advice, work, years, please      | 7.4                  |
| Other                | spy, volume, chart, support, bullish, low, trend, resistance, bearish, line        | 3.7                  |
| Other                | amet, calendar, releases, wed, link, thurs, tues, fri, analyst, close              | 0.5                  |
| Options / Risk       | option, spy, profit, strike, spread, risk, loss, value, selling, position          | 6.3                  |

Table 9: Topics Extracted from BTM Model